AD_____

Award Number: DAMD17-00-1-0405

TITLE: Identification of Genetic Modifiers of Breast Cancer Risk

PRINCIPAL INVESTIGATOR: Barbara L. Weber, M.D.

CONTRACTING ORGANIZATION: University of Pennsylvania
Philadelphia, PA 19104-3246

REPORT DATE: July 2003

TYPE OF REPORT: Final

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

DISTRIBUTION STATEMENT: Approved for Public Release;
Distribution Unlimited

The views, opinions and/or findings contained in this report are those of the author(s) and should not be construed as an official Department of the Army position, policy or decision unless so designated by other documentation.

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 074-0188

1. AGENCY USE ONLY (Leave blank)	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	July 2003	Final (1 Jul 2000 - 30 Jun 2003)

4. TITLE AND SUBTITLE
Identification of Genetic Modifiers of Breast Cancer Risk

5. FUNDING NUMBERS
DAMD17-00-1-0405

6. AUTHOR(S)
Barbara L. Weber, M.D.

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
University of Pennsylvania
Philadelphia, PA 19104-3246

E-Mail: weberb@mail.med.upenn.edu

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT (Maximum 200 Words)

The purpose of this study was to identify genetic modifiers of cancer risk in women with BRCA1 and BRCA2 mutations. We used two complementary strategies: 1) association studies in candidate genes from the immune surveillance and DNA damage response pathways and 2) a genome-wide scan using relative pairs with BRCA1 mutations to identify novel regions containing modifier genes. We have assembled a case-control sample set of 448 mutation carriers and a relative pairs set of 534 mutation carriers. We completed a sequencing survey of a panel of immune surveillance genes and determined the population frequency of the variants we identified. We examined a number of candidate genes and have data suggesting variants in TNF-α, IL-6, XPD and p53 may have a role in altering cancer risk in these high risk women. This work is important not only in leading to more refined cancer risk estimates for women with BRCA1 and BRCA2 mutations, but also will yield candidates for risk alleles in the general population and generate hypotheses for mechanisms that explain these effects. Once these mechanisms have been elucidated, these points in key pathways become excellent targets for preventative and therapeutic intervention.

14. SUBJECT TERMS
Cancer biology, etiology, BRCA1, susceptibility genes, low penetrance genes, linkage

15. NUMBER OF PAGES
53

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	Unlimited

Table of Contents

INTRODUCTION

The focus of this research study was the identification of genetic factors that influence cancer risk is women with BRCA1 and BRCA2 mutations. We collected DNA and information from a large retrospective cohort of women with BRCA1 mutations. These were assembled into two, overlapping study samples: 1) a case control set where all study samples were derived from women with BRCA1 mutations; cases were those with breast cancer and controls are those carriers that have not developed breast cancer and 2) a relative pairs set were all samples are matched with at least one family member who also had a BRCA1 mutation. These sets were used with two distinct methodologies to identify genetic modifiers of BRCA1 penetrance, including a candidate gene approach focused on a panel of genes involved in response to DNA damage and of genes important in modulating immune surveillance and a modified linkage approach to identify novel genes.

PROGRESS REPORT

Task 1: Screening of all genetic variants in a series of candidate genes (Months 1-18).

a. Collection of DNA samples from all collaborators.

This task was completed primarily during year 1. The sample set from which the case-control set for analysis of candidate genes was constructed includes 656 women with germline *BRCA1/2* mutations. These samples were ascertained in a retrospective fashion after identification of families with a history of breast and/or ovarian cancer at Creighton University, the Dana Farber Cancer Institute, The University of Michigan, Fox Chase Cancer Center, The University of Pennsylvania, The University of Utah, or Women's College Hospital (Toronto) between 1978 and 1997. The resulting case control sample of 448 women consists of 278 breast cancer cases and 170 matched controls.

The collection of relative pairs was completed in Year 2. We ascertained data on 600 BRCA1 mutation carriers. Of those, 242 female mutation carriers from 51 families had a relative with a known mutation and a DNA sample that could be included in the analysis. Ninety individuals from 21 families were used for the chromosome 5q linkage analysis and the remainder were used in year 3 for validation of this finding.

Subtasks b-f (PCR amplification of variant fragments and microsatellites (b), separation with automated sequencer (c), checking of automated data (d), data analysis (e) and reanalysis if indicated (f)) are completed sequentially for each gene undergoing analysis before conclusions can be drawn so they are considered as a single task with the following analyses.

<u>Immune surveillance genes</u>

We completed a comprehensive sequence analysis of 13 immune surveillance genes for the presence and frequency of 26 polymorphisms in a control population set in Year 1. This work has now been published and the manuscript is found in the Appendix (Martin, AM et al, 2003).

During year we arrayed the case-control samples into 96-well microtiter plates and began the analysis of a number of immune surveillance candidate genes (see below). We optimized PCR conditions for all polymorphisms on the ABI 3100 capillary sequencer and have screened 13 genes/26 polymorphisms (14 novel) in the case control set and have found the following:

- 7 polymorphisms where the variant allele is present at >15% frequency in both groups
  IL-1β,IL1-RN (2), IL-2, IL-6, IL-10, CTLA4
- 6 polymorphisms where the variant allele is present at < 15% frequency in only one group
  TNF-α (2), TNF-αR (2), IL-12p35, CTLA4
- 11 polymorphisms where the variant allele is present at <15% frequency in both groups
  TNF-α(4), IL-1α(2), IL-10, IL-12p35, IL-12p40(3), CTLA4

We have completing the statistical analysis of these polymorphisms in for associations with case status and age of diagnosis. We have evidence from these analyses that polymorphisms IL-6 may be associated with variable breast cancer risk in BRCA1/2 mutation carriers. These polymorphisms are now being analyzed in a validation set of an additional 500 mutation carriers (beyond the scope of this proposal but required for confirmation of the association).

DNA damage response genes

The analyses of XRCC1 and XPD (also called ERCC2) were completed in year 1. The genotypes examined included XRCC1 exon 6 Arg194Trp and exon 10 Arg399Gln and XPD exon 6 C>A, 156Arg, exon 10 Asp312Asn, exon 22 C>T, Asp711, and exon 23 Lys751Gln. Three of the four XPD polymorphisms showed statistically significant association with breast cancer risk in our case population. The Lys allele at Lys751Gln in exon 23 (age-adjusted OR: 1.89; 95% CI: 1.10-3.22), the C allele (C>T, Asp711) in exon 22 (OR: 2.02; CI: 1.11-3.66) and the C allele (C>A, 156Arg) in the exon 6 (OR: 3.96; CI: 1.92-816) showed association with increased breast cancer risk in BRCA1 mutation carriers. No association between genotypes and breast cancer risk was observed for the polymorphisms in the XRCC1 gene. These data are being confirmed in the validation set as described above.

The genetic polymorphisms and population frequencies for additional DNA damage response genes have now been made publicly available by Dr. Henry Mohrenweiser at Lawrence Livermore National Laboratory and as planned, we have completed our analyses of the genes that participate in the DNA repair-related BASC complex (which also includes BRCA1) (Table 1). We studied the following variants: ATM - 5'UTR 10805 A/G, D1853N, MLH1 - 5'UTR –93 G/A, MSH2 - IVS9 –9 T/C, IVS12 –6 T/C, and MSH6 - G36E. Of the 6 variants, only the glutamine variant at codon 36 of MSH6 may be associated with the diagnosis of breast cancer (OR=2.7, 95% CI 0.86-4.9) but in this data set the findings do not reach statistical significance. As BRCA1 and MSH6 function in different pathways of DNA damage repair, double strand break repair and mismatch repair respectively, it is possible that alterations in multiple pathways may be more important than multiple alterations in the same pathway.

Table 1. BASC Complex polymorphisms			
Gene	Variant	Reported Frequency	Our frequency
MLH1	%'UTR –93	50%	27%
MLH1	1219V/L A>G or C	13-34%	ND
MSH2	IVS9-9T>C	20%	26%
MSH2	IVS12-6T>C	23%	10%
MSH2	G36E G>A	25%	19%
ATM	5'UTR10805 A>G	28%	55%
ATM	S49C	0.5%	2%
ATM	P1054R	1.5%	3%
ATM	D1853N G>A	25%	16%

Task 2. Perform a genome wide search to identify regions that contain novel genes, which modify breast cancer risk in BRCA1 mutation carriers (Months 18-36).

a. PCR amplification of microsatellite markers at 10-15 cM intervals throughout the genome. A request for funding for this subtask was outside the scope of this proposal and was requested from both the Center for Inherited Disease Research and The Marshfield Center. Both proposals were turned down due to what was perceived as the highly speculative nature of the project. Despite this, we believe this is an important component of the search for modifier genes, as not all may be considered a priori as candidates. Thus we undertook a directed study of chromosomes 4 and 5q, as these regions are frequently lost in BRCA1-associated breast cancers. With a limited set of relative pairs we did not see evidence of an association with age of diagnosis on chromosome 4 but found very interesting evidence of linkage on chromosome 5q. These data were published during Year 2. Subtasks b(separation with automated sequencer gel apparatus), c (checking of all automated data analysis), d (submission of final data in linkage format output to Dr. Shugart) and e (reanalysis of samples as indicated by statistical analysis) are all described manuscript submitted with our 2002 Progress Report (Nathanson, KL et al, 2002) .

Task 3. Statistical analysis of data (Months 12-36)

a. Analysis of candidate gene variants using a cohort study design based on Cox proportional hazards models and a case control design based on logistic regression analysis. We evaluated the relationship between genotypes at one of our high-priority candidate genes, IL-6, and breast or ovarian cancer using a nested case-control analysis approach. Overall, we saw no association of breast cancer with IL-6 genotypes carrying a G allele at position -174 (OR=1.0, 95% CI: 0.5-1.8). Similarly, we saw no effect of these genotypes on breast cancer risk when stratified by smoking, parity, or oral contraceptive use. Similarly, we saw no effect of IL-6 -174 GG genotype with ovarian cancer overall (OR=0.5, 95% CI: 0.2-1.8), and no effect of genotype on ovarian cancer risk when stratified by talc use, parity, or oral contraceptive use. However, we did observed a significant reduction of ovarian

cancer risk by IL-6 genotypes among ever smokers (OR=0.14, 95% CI:0.04-0.96; OR adjusted for parity, year of birth, and oral contraceptive use) but not in never smokers (OR=1.62, 95% CI: 0.19-13.71). This observation suggests that inflammatory responses to cigarette smoking, possibly mediated by IL-6, may affect ovarian cancer risks. These observations need to be further evaluated in a larger sample set considering dose and duration of cigarette smoking in addition to other ovarian cancer risk factors. Analyses of additional immune surveillance and DNA repair genotypes are currently underway.

b. Analysis of linkage data using both model-based and model-free approaches. We will use both identity-by-state and identity-by-descent methods, including APM and SimIBD. We hypothesized that the modifier genes might be located in regions of allelic imbalance in the tumors of *BRCA1* mutation carriers, as have been reported on chromosomes 4p, 4q, and 5q. In order to determine whether novel genetic modifiers of *BRCA1*–associated breast cancer penetrance in these regions exist, we used non-parametric linkage analysis methods to determine whether specific chromosome 4p, 4q and 5q haplotypes were observed preferentially in breast cancer cases among women with *BRCA1* mutations. No significant linkage on chromosome 4p or 4q was observed associated with breast cancer risk in BRCA1 mutation carriers. However, we observed a significant linkage signal at D5S1471 on chromosome 5q (p= 0.009) in all the families analyzed together. The significance of this observation increased in the subset of families with an average of breast cancer diagnosis less than 45 years (p=0.003). These results suggest the presence of one or more genes on chromosome 5q33-34 that modify breast cancer risk in *BRCA1* mutation carriers. The approach described here may be utilized to identify penetrance modifiers in other autosomal dominant syndromes. This work was completed and published in May 2002. Since that time we have continued the analysis of this region using a case-control analysis of candidate genes in the region. At present, preliminary data suggest that a strong candidate has been identified. This work is being prepared for publication,

These steps are described as part of Tasks 1 and 2 and have been completed for all the genes and chromosomal regions described in those analyses.

KEY RESEARCH ACCOMPLISHMENTS

- Collection of matched case-control set of 448 BRCA1 and BRCA2 mutation carriers (Year1)
- Collection of a relative pair sample set containing data on 600 BRCA1 mutation carriers and DNA samples on 242 mutation carriers as components of a relative pair (Year 2)
- Completion of a comprehensive sequencing survey of immune surveillance genes for polymorphic variants (Year 1)

- Analysis of the population frequency of 26 polymorphisms in 13 immune surveillance genes (Year 1)

- Genotyping of 26 polymorphisms in 13 immune surveillance genes in the BRCA1/BRCA2 mutation carrier case-control set (Year 2)

- Identification of the immune surveillance genes TNF-α and IL-6 as candidate risk modifiers in the set of BRCA1/2 mutation carriers (Year 2)

- Exclusion of the immune surveillance genes IL1-α and -β, IL1-RN, IL-2, IL-10, CTLA4, TNF-αR, IL12p35 and IL12p49 from further analysis as candidate modifier genes in this set (Year 2)

- Analysis of the population frequency of nine polymorphisms in four DNA damage response gene polymorphisms (Year 2)

- Analysis of associated haplotypes in ATM to facilitate polymorphism typing (Year 3)

- Genotyping of 16 polymorphisms in seven DNA damage response genes in the BRCA1/BRCA2 mutation carrier case-control set (Years 1 and 2)

- Identification of the DNA damage response genes TP53 and XPD (Year 1) and MSH6 (Year 2) as a candidate risk modifiers in the set of BRCA1/2 mutation carriers

- Exclusion of XRCC1 (Year 1) and ATM, MLH1 and MSH2 (Year 2) from further analysis as a risk modifier in this set of BRCA1/2 mutation carriers

- Genotyping of 25 polymorphic microsatellite repeats on chromosomes 4 and 5q at an average of 12 cM intervals in 73 BRCA1 mutation carrier relative pairs (Year 2)

- Exclusion of chromosome 4 as a locus for candidate modifier genes in this sample set (Year 2)

- Identification of chromosome 5q as a candidate region to contain a modifier of BRCA1-related breast cancer penetrance, with a maximum likelihood score at locus D5S1471 (Year 2)

- Extended analysis of IL-6 as a modifier gene – analyses attached, manuscript under preparation (Year 3).

- Analysis and publication of results from the chr 4 and 5 linkage scan for modifiers (Years 2 and 3)

- Screening and analysis of candidate genes in the chr 5 linkage region for polymorphisms association with altered risk (Year 3) .

REPORTABLE OUTCOMES

- A manuscript describing the immune surveillance gene polymorphism discovery and frequency evaluation has been published. (manuscript attached).

Martin AM, Athanasiadis G, Greshock JD, Fisher J, Lux MP, Calzone K, Rebbeck TR, Weber BL. Population frequencies of single nucleotide polymorphisms (SNPs) in immuno-modulatory genes. Hum Hered. 2003;55(4):171-8.

- A manuscript is under preparation describing the XPD and XRCC1 analyses. This work was presented in abstract form in at AACR in March, 2001 and an updated, expanded analysis was presented at ASHG in October, 2001.

B. Amirimani, S.L. Neuhausen, T. Tran T.R. Rebbeck, and B.L. Weber.Polymorphisms in XRCC1 and XPD as Breast Cancer Risk Modifiers in BRCA1 Mutation Carriers. Proceeding, American Association for Cancer Research, 2001.

B. Amirimani, S.L. Neuhausen, T. Tran T.R. Rebbeck, and B.L. Weber.Polymorphisms in XRCC1 and XPD as Breast Cancer Risk Modifiers in BRCA1 Mutation Carriers. Proceedings, American Society of Human Genetics, October 2001.

- An abstract describing IL-6 and TNF a as candidate modifiers of BRCA1 penetrance was presented at ASHG in October, 2001.
-

A-M. Martin, P.A. Kanetsky, G. Athanasiadis, J.D. Greshock, T.R. Rebbeck, B.L. Weber Immune surveillance genes and breast cancer: do *IL-6* or *TNFa* modify *BRCA1* penetrance? Proceedings, American Society of Human Genetics, October 2001.

- An abstract describing the analysis of the BASC complex genes as modifiers of BRCA1 penetrance was presented at ASHG in October, 2001.

KL Nathanson, R Letrero, P Kanetsky, Romaruddin, TR Rebbeck, BL Weber. Variants in the genes that encode the BRCA1-associated genome surveillance complex (BASC) in BRCA1 mutation carriers, Proceedings, American Society of Human Genetics, October 2001

- A manuscript describing the p53 polymorphism effect in women with BRCA1/2 mutations and multiple primary cancers has been published (manuscript attached).

Martin, A.M., Kanetsky, P.A., Amirimani, B., Colligon, T.A., Athanasiadis, G., Shih, H., Gerrero,M.R., Calzone, K.A., Rebbeck, T.R., Weber, B.L. Germline TP53 mutations in breast cancer families with multiple primary cancers. J Med Genet 40:e4, 2003.

- A manuscript has been published describing the analysis of chromosomes 4 and 5q for candidate modifier loci.

Nathanson, K.L., Shugart, Y.Y., Omaruddin, R., Szabo, C., Golgar, D., Rebbeck, T., Weber, B.L. CGH-target linkage analysis reveals a possible BRCA1 modifier locus on chromosome 5q. Human Molec Genet, 11:1327-1132, 2002.

- A manuscript that reconciles several previously published ATM haplotypes has been published.

Letrero R, Weber BL, Nathanson KL. Resolving ATM haplotypes in whites. Am J Hum Genet. 2003 Apr;72(4):1071-3.

- Patents and/or licenses: None.

- Degrees obtained: None.

- Repositories, data banks and informatics tools: No new ones have been created – this work is being performed retrospectively.

- Funding applied for on the basis of this work: Dr. Nathanson, the postdoctoral fellow who performed the portion of this work aimed at evaluating DNA damage response genes has applied for an NIH RO1 to continue this work. That grant is currently under review.

- Employment/research opportunities: One postdoctoral fellow (B. Amirimani) has completed her training with the analysis of the XRCC1 and XPD analyses and has obtained permanent employment based on this work (Year 1). She does not working on this project in her new position. A second postdoctoral fellow (A-M. Martin) has completed her training with the work on the immune surveillance genes. She has obtained a full time faculty position in an affiliated hospital. She continued to supervise work on the immune surveillance gene polymorphisms until that portion of the project was completed. A third postdoctoral fellow (K.L. Nathanson) completed her fellowship and obtained a faculty position at the University of Pennsylvania in the Department of Medicine. This work has formed the basis for her independent career.

CONCLUSIONS

This work supports the existence of multiple genetic modifiers of BRCA1/2-related breast cancer penetrance. We have evidence that genetic variants in TNF-α, IL-6, p53 and XPD and MSH 6 may function in this capacity. In addition, we have evidence for a candidate locus on chromosome 5q based on a modified linkage approach. This work is important not only in ultimately leading to more refined cancer risk estimates for women with BRCA1 and BRCA2 mutations, but will also yield candidates for risk alleles in the general population as well as generate hypotheses for mechanisms that explain these

effects. Once these mechanisms have been elucidated, these points in key pathways become excellent targets for preventative and therapeutic intervention.

REFERENCES

None

APPENDIX

Manuscripts as described in Reportable Outcomes
IL-6 Analysis Part 1
IL-6 Analysis Part 2

ELECTRONIC LETTER

Germline *TP53* mutations in breast cancer families with multiple primary cancers: is *TP53* a modifier of *BRCA1*?

A-M Martin, P A Kanetsky, B Amirimani, T A Colligon, G Athanasiadis, H A Shih, M R Gerrero, K Calzone, T R Rebbeck, B L Weber

J Med Genet 2003;**40**:e34(http://www.jmedgenet.com/cgi/content/full/40/4/e34)

Somatic mutations in *TP53* are the most frequent events in human cancer and lead to inactivation of the gene, loss of tumour suppressor function, and in some cases generation of a dominant negative form of p53.[1–3] Eleven exons make up the primary transcript of *TP53*, of which exons 2-11 encode the protein. Five conserved domains exist in exons 1, 4, 5, 7, and 8,[4] which are considered essential for normal p53 function. Approximately 90% of disease associated mutations occur in these domains, with mutations in five codons (175, 245, 248, 249, and 273) accounting for approximately 20% of all mutations reported to date.

Germline mutations in *TP53* cause Li-Fraumeni syndrome (LFS), a familial association of childhood leukaemia, brain cancer, soft tissue sarcoma, and adrenal cortical carcinoma,[5 6] as well as other cancers such as breast cancer, melanoma, germ cell tumours, and carcinomas of the lung, pancreas, and prostate.[7 8] Cancers characteristically develop at unusually early ages and multiple primary tumours are frequent. Susceptibility to cancer in these families follows an autosomal dominant pattern of inheritance[7] and among families with a known germline *TP53* mutation the probability of developing any invasive cancer (excluding carcinomas of the skin) approaches 50% by the age of 30, compared to an age adjusted population incidence of cancer of 1%. It is estimated that more than 90% of *TP53* mutation carriers will develop cancer by the age of 70.[9]

In addition to the numerous mutations, *TP53* also contains several polymorphisms that may alter its activity. In particular, at nucleotide 215 (codon 72) there is a single base pair variant (g.215G>C) in the coding region, which results in a substitution of proline for arginine in the protein sequence.[10] The frequency of this polymorphism varies from 26-35%[11–13] and it appears to affect protein function. The R72 variant of TP53 is believed to be more sensitive to human papillomavirus (HPV) induced degradation by the E6 oncoprotein than the 72P variant, and is thought to be of functional significance in HPV associated tumours[14] such as cervical tumours.[15–17] Furthermore, some, but not all studies document an over-representation of R72 variant in cervical cancer patients compared to a control population.[18 19] However, other reports suggest the association of the 72P variant with incidence of squamous cell carcinoma of the head and neck[20] and lung adenocarcinomas in smokers.[21]

In families with multiple cases of breast cancer that do not fit the criteria for LFS, the frequency of *TP53* germline mutations has been investigated in multiple studies,[22–28] documenting that *TP53* mutations account for <1% of site specific breast cancer families.[23] However, among LFS families, there is a very high incidence of early onset breast cancer. Taken together, these data suggest that germline *TP53* mutations are strongly associated with hereditary breast cancer susceptibility but almost exclusively in the context of LFS.

Because of the high penetrance of early onset breast cancer and the known increased incidence of multiple primary cancers in LFS families (50% by 30 years of age),[9] we investigated whether deleterious germline mutations in *TP53* and/or the R72P polymorphism were associated with multiple primary cancers (in which one was breast cancer) in families with ≥2 breast cancers but no evidence of LFS. One previous study investigated the frequency of germline *TP53* mutations with bilateral breast cancer[29] and found no *TP53* mutations; however, only 19 samples were tested. In the current study we determined the frequency of deleterious *TP53* germline mutations in 172 breast cancer families, with and without multiple primary cancers. Germline *BRCA1* or *BRCA2* mutation status was known in all subjects[30]; 43% of women with multiple primary cancers and 12% of women with breast cancer only had

Key points

- Eighty eight women with breast cancer and a personal or family history of multiple primary cancers (MPC) (including ovarian cancer) and 84 women with a personal and family history of breast cancer only (BC) were studied. All women had been previously screened for germline *BRCA1* and *BRCA2* mutations; 38 (43%) of MPC women and 10 (12%) of BC women had a mutation in one of these two genes.
- We determined the frequency of deleterious germline *TP53* mutations, as well as the common R72P polymorphism in *TP53* and investigated the association of this polymorphism with the development of cancers in the entire study set. We also evaluated the association between R72P and breast cancer penetrance in the subset of women with known *BRCA1* or *BRCA2* mutations.
- One woman, from a family with breast cancer only, was found to have a deleterious *TP53* mutation (exon 7, G245S); no deleterious *TP53* mutations were detected in the families with cases of multiple primary cancers. The common R72P polymorphism was seen at a frequency of 41% in the entire sample. MPC women were more likely to be homozygous for R72 compared to BC women (p=0.05, OR 2.83, 95% CI 1.2 to 6.9), an association that was more striking in women with a *BRCA1* or *BRCA2* mutation (OR 6.1, 95% CI 1.4 to 26.4).
- We also found that the presence of a 72P allele was associated with an earlier age of breast cancer diagnosis among *BRCA1* mutation carriers (p=0.05), suggesting that the R72P polymorphism may be a modifier of *BRCA1* penetrance.

Abbreviations: MPC, multiple primary cancers; BC, breast cancer only; LFS, Li-Fraumeni syndrome; HPV, human papillomavirus

Table 1 Non-breast primary cancers in families with multiple primary cancers

Cancer	No of families with cancer*
Ovarian	29
Colorectal	13
Non-melanoma skin	11
Thyroid	6
Endometrial	6
Cervix	5
Leukaemia	4
Lymphoma	3
Others†	13

*Five patients had two or more non-breast cancers, so the number of cancers does not equal the number of patients.
†Other primary cancers include melanoma, brain, head/neck, sarcoma, lung, kidney, and pharynx.

either a *BRCA1* or *BRCA2* mutation. In addition, we established the frequency of the common exon 4 polymorphism (R72P) in this sample and evaluated whether this polymorphism may be a modifier of breast cancer penetrance in the presence of *BRCA1* or *BRCA2* mutations.

MATERIALS AND METHODS

Patient population

All families were recruited from clinics at the University of Michigan (1993-1995) and the University of Pennsylvania (1995-1998). Patients were either self- or physician referred because of a perceived risk of inherited susceptibility to breast cancer. All women consented to genetic testing for clinical and/or research purposes. Personal and family histories of all cancers were recorded, including age of diagnosis of all cancers and the number of related women in each family at risk for breast cancer (age ≥20 years). Pathology reports were obtained on all probands and on other family members when possible. The testing protocol was approved by duly constituted institutional review boards at both the University of Michigan and the University of Pennsylvania.

Eighty-eight women were from families with at least two cases of breast cancer and at least one woman affected with both a primary breast cancer and a primary non-breast cancer (denoted MPC). Eighty-four (95%) MPC women had two primary cancers, and four MPC women (5%) had three or more primary cancers. All non-breast malignancies were considered, including non-melanoma skin cancers. An additional 84 women were from families with at least two cases of breast

Table 2 Description of relatives providing sample for testing

Sample number tested (relative)	Age of breast cancer in relative	Average age of diagnosis of additional cancer(s) in multiply affected subject	Second primary cancer in multiply affected subject	Relationship of subject providing DNA for testing
229	30	58	Ovarian	Daughter
522	56	63	Ovarian	Niece
743	69	74	BCC, cervical NHL	Niece
673	40	43	BCC	Daughter
513	38	30	Cervical	Daughter
641	44	30	Colon	Cousin
813	36	44	Ovarian	Niece
852	56	53	BCC, endometrial	Niece
834	45	40	Cervical	Sister
1019	50	63	Ovarian	Daughter
842	27	47	Ovarian	Sister
975	57	50	BCC	Sister
1965	57	25	Thyroid	Daughter
910	74	55	Thyroid	Daughter
1946	44	84	Leukaemia	Daughter
1901	30	65	Lung	Niece
1907	46	86	Colon, rectal	Daughter
1208	41	51	Ovarian	Sister
1708	72	69	Skin, cervical, endometrial	Sister
1785	51	49	Colon	Granddaughter
1844	52	67	Colon	Daughter
1773	50	51	Cervical	Niece
1762	45	74	Colon	Daughter
1797	36	50	Ovarian	Daughter
1748	42	72	Hodgkin's lymphoma	Granddaughter
1722	50	83	Endometrial	Granddaughter
1719	40	23	Ovarian	Sister
1718	35	36	Ovarian	Granddaughter
1763	41	56	Colon	Daughter
1794	40	40	Ovarian	Cousin
1783	53	53	Throat	Double cousin
1853	57	57	Brain	Grandniece
320	40	40	Pituitary	Niece
1987	55	55	Ovarian	Daughter
1909	70	70	Leukaemia	Granddaughter
1954	61	61	Thyroid, colon	Sister
1851	43	74	Leukaemia	Cousin
1806	49	49	Melanoma	Daughter
1863	39	39	Ovarian	Daughter
1993	67	67	Colon	Cousin
2074	95	95	Colon	Niece
2038	50	52	Ovarian	Daughter
2067	50	50	BCC	Daughter
2233	41	42	ALL	Niece
2241	61	52	Ovarian	Cousin

Table 3 Primer sequences

Exon	Primer sequences	Annealing temp (°C)	Reference
2/3	Forward: 5′-ggatccccactttcctctt-3′ Reverse: 5′-agcatcaaatcatccattgc-3′	57	
4	Forward: 5′-gacctggtcctctgactgct-3′ Reverse: 5′-atacggccaggcattgaagt-3′	54	
5/6	Forward: 5′-tgccctgactttcaactctgt-3′ Reverse: 5′-ttaaccctcctcccagaga-3′	54	
7	Forward: 5′-gccacaggtctccccaagg-3′ Reverse: 5′-aggggtcagcggcaagcaga-3′	55	Evans *et al*[33]
8/9	Forward: 5′-caaggggtggttgggagtaga-3′ Reverse: 5′-acttgataagagggtcccaag-3′	54	
10	Forward: 5′-atgttgcttttgatccgtca-3′ Reverse: 5′-ctttccaacctaggaaggca-3′	54	
11	Forward: 5′-agccacctgaactcaaaaa-3′ Reverse: 5′-aatggcaggggagggagaga-3′	55	Evans *et al*[33]

cancer but no cases of multiple primary cancers (denoted BC). DNA was available from at least one woman with multiple primary cancers in 43 families. In the remaining 45 multiple primary cancer families, the multiply affected woman was dead (n=33) or unavailable (n=12). In these families, *TP53* screening was undertaken using DNA from the closest female relative diagnosed with breast cancer. Table 1 provides a description of the cancers reported in subjects with multiple primary cancers and table 2 is a detailed description of the female relatives of a multiply affected woman, who provided samples for testing. All samples were previously screened for germline mutations in *BRCA1* and *BRCA2*[30]; 38 MPC women and 10 BC women had a *BRCA1* or *BRCA2* germline mutation.[30]

Mutation analysis

DNA was extracted from peripheral blood mononuclear cells and stored in TE at 4°C. The entire 10 exon coding domain and flanking splice site regions of *TP53* were amplified using seven PCR primer sets (table 3). PCR amplification was performed in a final volume of 20 µl containing 80 ng of DNA, 1.5 mmol/l $MgCl_2$, 10 mmol/l Tris-HCl (pH 8.3), 50 mmol/l KCl, 0.2 mmol/l each of dCTP, dATP, dTTP, dGTP (Amersham Pharmacia Biotech), each primer at 1.0 µmol/l, and 1.0 unit of *Taq* polymerase (Boerhinger Manheim). Annealing temperatures were optimised for each primer set and ranged from 55-60°C. Variants were identified by conformation sensitive gel electrophoresis (CSGE) as previously described[30] and characterised by direct sequencing using the ABI Prism 377 after reamplification from source DNA. All mutation nomenclature is reported using the recommendations of den Dunnen and Antonarakis.[31]

Statistical analysis

Differences in *TP53* mutation frequency between the MPC and BC groups were assessed using χ^2 analysis. Odds ratios (OR) and 95%t confidence intervals (95% CI) were reported. In those instances where expected cell counts fell below five, we used exact methods to determine the 95% CI.[32] Furthermore, we used the Mann-Whitney U-Wilcoxon rank sum test to determine whether *TP53* genotypes altered the median age of first breast cancer diagnosis within categories defined by *BRCA1* or *BRCA2* mutation status.

RESULTS

DNA from one woman from a BC family and no *BRCA1* or *BRCA2* mutation showed an abnormal CSGE profile in *TP53* exon 7. Sequence analysis of this variant showed a G to A transition at the first nucleotide in codon 245 resulting in a glycine-serine change at this position (G245S). No presumed deleterious *TP53* mutations were seen in the MPC group (MPC=0%, BC=1.2%, p=0.31).

The proline allele of the R72P polymorphism was seen at a frequency of 41% in the entire sample. The distribution of R72P genotypes within groups is presented in table 4. Owing to the small number of homozygous 72P genotypes, all 72P alleles were combined into one group (P/*). When we performed subgroup analyses in women with *BRCA1* or *BRCA2* mutations, MPC women were six times more likely to have the homozygous R72 genotype than BC women (OR=6.1, 95% CI 1.4 to 26.4) (table 5). However, because of the small sample size, a statistically significant association between the homozygous R72 genotype and MPC could not be confirmed separately in an analysis of only *BRCA1* mutation carriers or only *BRCA2* mutation carriers.

Table 4 *TP53* exon 4 R72P genotypes in subjects with multiple primary cancers (MPC) compared to subjects with breast cancer only (BC)

TP53 genotype	MPC (n=43) No	(%)	BC (n=84) No	(%)	
R/R	35	(81)	51	(61)	
R/P	6	(14)	28	(33)	
P/P	2	(5)	5	(6)	
P/	8	(19)	33	(39)	OR=2.83 (1.2, 6.9)

P/ is the combined genotypes of R/P and P/P used for statistical analysis owing to the rarity of the P allele.

Table 5 *TP53* R72P genotypes in MPC and BC families by *BRCA1/2* mutation status

TP53 genotype	MPC (n=43) No	(%)	BC (n=84) No	(%)	OR 95% CI
BRCA1 or *BRCA2* mutation					
R/R	19	(83)	7	(44)	6.1 (1.4, 26.4)
P/*	4	(17)	9	(56)	1.0
BRCA1 mutation					
R/R	16	(80)	3	(43)	5.3 (0.83, 34.1)
P/*	4	(20)	4	(57)	1.0
BRCA2 mutation					
R/R	5	(83)	4	(44)	6.3 (0.50, 77.5)
P/*	1	(17)	5	(56)	1.0
No detectable mutation					
R/R	16	(80)	44	(65)	2.2 (0.66, 7.3)
P/*	4	(20)	24	(35)	1.0

P/* = R72P or P72P.
Three women in the sample set had both a *BRCA1* and a *BRCA2* mutation.

Table 6 *TP53* R72P genotypes by age of diagnosis of breast cancer and *BRCA1/2* mutation status

TP53 genotype	Median age	(IQR)	p value
BRCA1 or *BRCA2* mutation			
R/R (n=26)	42	(35-51)	0.01
P/* (n=13)	32	(30-38)	
BRCA1 mutation			
R/R (n=19)	46	(35-51)	0.05
P/* (n=8)	32	(30-41.5)	
BRCA2 mutation			
R/R (n=9)	39	(35-46)	0.39
P/* (n=6)	35	(30-43)	
No detectable mutation			
R/R (n=60)	49	(37-61.5)	0.98
P/* (n=28)	50	(42.5-57)	

Three women in the sample set had both a *BRCA1* and a *BRCA2* mutation.

In an evaluation of the R72P polymorphism as a modifier of breast cancer penetrance in women with germline *BRCA1* or *BRCA2* mutations, we found that in the combined *BRCA1* and *BRCA2* mutation carrier analysis, the presence of any 72P allele was associated with an earlier median age of breast cancer diagnosis (median age=32, interquartile range (IQR) 30-38) compared with the homozygous R72 genotype (median age=42, IQR 35-51, $p<0.01$) (table 6). This association was limited to *BRCA1* mutation carriers (median age=32, IQR 30-41.5, $p<0.05$) and was not seen in *BRCA2* mutation carriers (median age=35, IQR 30-43, $p<0.39$) (table 6).

DISCUSSION

In this study, we screened all 10 coding exons of *TP53* in women with a personal history of breast cancer with or without a personal or family history of multiple primary cancers. These women previously had been characterised for *BRCA1* and *BRCA2* mutations.[30] We identified one potential deleterious missense mutation (G245S) in a member of a family with a history of site specific breast cancer only. This patient did not carry a germline mutation in either *BRCA1* or *BRCA2*. The G245S missense mutation has been reported previously in the germline of a woman with breast cancer[31] and the germline of a man with sarcoma.[32] Our proband was diagnosed with breast cancer at the age of 29. In addition, her sister was diagnosed with breast cancer at the age of 27 and went on to develop a second primary breast cancer at 31 years of age (fig 1). However, the G245S mutation was not detected in DNA from the sister's first breast tumour. In addition, there was no allelic loss of flanking *TP53* in that tumour (data not shown). Thus it is possible that either the proband's sister is a phenocopy or the *TP53* mutation is not the relevant source of breast cancer susceptibility in this family, but this would need to be confirmed by testing the germline DNA, which was unavailable. Thus, we conclude that germline *TP53* mutations are not an important cause of multiple primary cancers outside the setting of LFS.

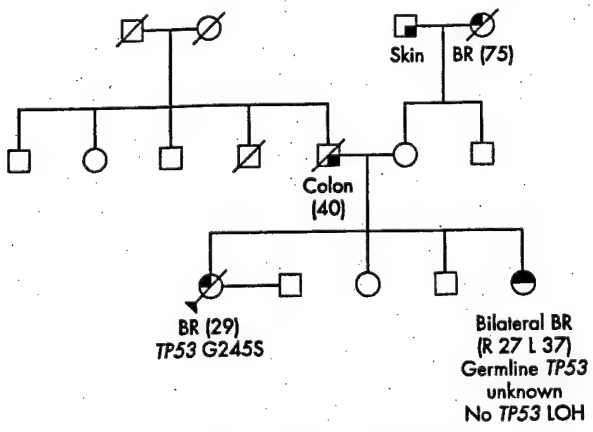

Figure 1 Pedigree of a patient with germline *TP53* mutation. Numbers in parentheses indicate age of cancer diagnosis.

In this sample set the 72P allele was found at a frequency of 41%, somewhat higher than the previously reported 26-35%.[11-13] Nonetheless, we observed a six-fold higher frequency of the homozygous R72 genotype among MPC women with *BRCA1* or *BRCA2* mutations compared to BC women with a *BRCA1* or *BRCA2* mutation. These data suggest that women who are homozygous for the R72 allele and have a mutation in *BRCA1* or *BRCA2* may be at increased risk for developing multiple primary cancers. Although contrary to the study by Brose *et al*,[34] who showed that *BRCA1* mutations do not confer an increased risk of most additional primary cancers, nonetheless this present study did not explore the combined association of both a *BRCA1* mutation and the *TP53* R72P polymorphism and its potential role as a modifier of *BRCA1* associated breast cancer risk.

Other studies associating the homozygous R72 allele and increased cancer risk have been reported. In the first such example, the association between *TP53* polymorphisms and human papillomavirus (HPV) associated cervical cancer was examined, suggesting that women who were homozygous for the R72 allele were seven times more susceptible to HPV related cervical cancer than with at least one 72P allele.[35] However, these data have been difficult to replicate and an equal number of studies have either confirmed[36-38] or disputed[39-43] the R72 association with cervical cancer.

In the subset analysis of the R72P polymorphism as a candidate modifier of breast cancer penetrance in *BRCA1/2* mutation carriers, we observed that the presence of a 72P allele was associated with an earlier age of breast cancer diagnosis among women with a *BRCA1* mutation. One possible explanation for the association of 72P with earlier onset breast cancer in *BRCA1* mutation carriers and R72 with MPC would be excess or earlier mortality among women with an earlier age of diagnosis of breast cancer (that is, those with the 72P). Thus if women homozygous for R72 may live longer, they may have a greater likelihood of developing a second cancer.

BRCA1 physically interacts with p53 in vitro and both BRCA1 and BRCA2 physically interact with p53 in vivo resulting in enhanced p53 mediated transcription.[44-46] There are two p53 binding sites in BRCA1; one is close to the nuclear localisation signal in the N-terminal region of exon 11[45] and one is in the most C-terminal BRCT domain.[47] Deletion of the N-terminal exon 11 p53 binding site prevents in vitro interaction of the two proteins and abrogates the coactivation effect of BRCA1 on p53 responsive promoters such as bax, p21, and GADD45.[45 48] In addition, a truncation mutant of BRCA1 that retains the p53 interacting site but removes the C-terminal BRCA1 transactivation domain acts as a dominant inhibitor of p53 dependent transcription.[45] Finally, *TP53* mutations are more common in *BRCA1* associated breast cancers than sporadic or *BRCA2* associated tumours. Somatic *TP53* mutations have been reported in as many as 80% of *BRCA1* associated tumours,[49 50] leading to the speculation that *TP53* mutations, or another component of the relevant pathway, maybe required before *BRCA1* related tumorigenesis can proceed.[31] Recent data from murine models strongly support this hypothesis.[52]

Our data provide additional support for a critical role of the p53/BRCA1 interaction in tumorigenesis, suggesting an association between *TP53* variants and cancer risk in women with *BRCA1* mutations. Thus, it is possible that the R72P polymorphism in *TP53* subtly alters the p53/BRCA1 interaction and in turn alters BRCA1 associated tumorigenesis.

In summary, we provide evidence that germline mutations in *TP53* are rarely associated with the presence of multiple

primary cancers in breast cancer families and support previous studies suggesting that *TP53* mutations account for less than 1% of hereditary susceptibility to breast cancer. However, we found presence of the homozygous R72 allele was associated with a six-fold increased risk for the development of multiple primary cancers among subjects with a germline *BRCA1* or *BRCA2* mutation. Finally, we provide preliminary evidence that the arginine allele of R72P in exon 4 of *TP53* may modify *BRCA1* associated breast cancer risk, using age of diagnosis as a surrogate for penetrance.

ACKNOWLEDGEMENTS

This work was supported by grants from the Susan G Komen Breast Cancer Foundation (PDF99-003024 to A-MM) and a grant from the NCI (CA 76417) to BLW.

Authors' affiliations

A-M Martin, Laboratory of Molecular Pathology, Department of Pathology, Pennsylvania Hospital, Philadelphia, USA
A-M Martin, B Amirimani, T A Colligon, G Athanasiadis, H A Shih, M R Gerrero, B L Weber, Department of Medicine, University of Pennsylvania, Philadelphia, USA
A-M Martin, T A Colligon, G Athanasiadis, M R Gerrero, K Calzone, B L Weber, Abramson Family Cancer Research Institute, University of Pennsylvania Cancer Center, Philadelphia, USA
P A Kanetsky, T R Rebbeck, Department of Biostatistics and Epidemiology, University of Pennsylvania, Philadelphia, USA
P A Kanetsky, T R Rebbeck Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, USA

Correspondence to: Dr B L Weber, Room 514, BRB II/III, 421 Curie Boulevard, Philadelphia, PA 19104, USA; weberb@mail.med.upenn.edu

REFERENCES

1 **Kern SE**, Pietenpol JA, Thiagalingam S, Seymour A, Kinzler KW, Vogelstein B. Oncogenic forms of p53 inhibit p53-regulated gene expression. *Science* 1992;**256**:827-30.
2 **Vogelstein B**, Kinzler KW. p53 function and dysfunction. Cell 1992;**70**:523-6.
3 **Marutani M**, Tonoki H, Tada M, Takahashi M, Kashiwazaki H, Hida Y, Hamada J, Asaka M, Moriuchi T. Dominant-negative mutations of the tumor suppressor p53 relating to early onset of glioblastoma multiforme. *Cancer Res* 1999;**59**:4765-9.
4 **Ko LJ**, Prives C. p53: puzzle and paradigm. *Genes Dev* 1996;**10**:1054-72.
5 **Li FP**, Fraumeni JF. Soft-tissue sarcomas, breast cancer, and other neoplasms: familial syndrome? *Ann Intern Med* 1969;**71**:747.
6 **Li FP**, Fraumeni JF. Prospective study of a family cancer syndrome. *JAMA* 1982;**247**:2692-4.
7 **Strong LC**, Stine M, Norsted TL. Cancer in survivors of childhood soft tissue sarcoma and their relatives. *J Natl Cancer Inst* 1987;**79**:1213-20.
8 **Li FP**, Fraumeni JF, Mulvihill JJ. A cancer family syndrome in 24 kindreds. *Cancer Res* 1988;**48**:5358.
9 **Malkin D**, Li FP, Strong LC, Fraumeni JF Jr, Nelson CE, Kim DH, Kassel J, Gryka MA, Bischoff FZ, Tainsky MA, *et al.* Germ line p53 mutations in a familial syndrome of breast cancer, sarcomas, and other neoplasms. *Science* 1990;**250**:1233-8.
10 **Harris N**, Brill E, Shohat O, Prokocimer M, Wolf D, Arai N, Rotter V. Molecular basis for heterogeneity of the human p53 protein. *Mol Cell Biol* 1986;**6**:4650-6.
11 **Kawajiri K**, Nakachi K, Imai K, Watanabe J, Hayashi S. Germ line polymorphisms of p53 and CYP1A1 genes involved in human lung cancer. *Carcinogenesis* 1993;**14**:1085-9.
12 **Sjalander A**, Birgander R, Hallmans G, Cajander S, Lenner P, Athlin L, Beckman G, Beckman L. p53 polymorphisms and haplotypes in breast cancer. *Carcinogenesis* 1996;**17**:1313-16.
13 **Wang-Gohrke S**, Rebbeck TR, Besenfelder W, Kreienberg R, Runnebaum IB. p53 germline polymorphisms are associated with an increased risk for breast cancer in German women. *Anticancer Res* 1998;**18**:2095-9.
14 **Soussi T**, Beroud C. Assessing TP53 status in human tumours to evaluate clinical outcome. *Nat Rev Cancer* 2001;**1**:233-40.
15 **Kammer C**, Tommasino M, Syrjanen S, Delius H, Hebling U, Warthorst U, Pfister H, Zehbe I. Variants of the long control region and the E6 oncogene in European human papillomavirus type 16 isolates: implications for cervical disease. *Br J Cancer* 2002;**86**:269-73.
16 **Zehbe I**, Tachezy R, Mytilineos J, Voglino G, Mikyskova I, Delius H, Marongiu A, Gissmann L, Wilander E, Tommasino M. Human papillomavirus 16 E6 polymorphisms in cervical lesions from different European populations and their correlation with human leukocyte antigen class II haplotypes. *Int J Cancer* 2001;**94**:711-16.
17 **Zehbe I**, Wilander E, Delius H, Tommasino M. Human papillomavirus 16 E6 variants are more prevalent in invasive cervical carcinoma than the prototype. *Cancer Res* 1998;**58**:829-33.
18 **Zehbe I**, Voglino G, Wilander E, Genta F, Tommasino M. Codon 72 polymorphism of p53 and its association with cervical cancer. *Lancet* 1999;**354**:218-19.
19 **Zehbe I**, Voglino G, Wilander E, Delius H, Marongiu A, Edler L, Klimek F, Andersson S, Tommasino M. p53 codon 72 polymorphism and various human papillomavirus 16 E6 genotypes are risk factors for cervical cancer development. *Cancer Res* 2001;**61**:608-11.
20 **Shen H**, Zheng Y, Sturgis EM, Spitz MR, Wei Q. P53 codon 72 polymorphism and risk of squamous cell carcinoma of the head and neck: a case-control study. *Cancer Lett* 2002;**183**:123-30.
21 **Fan R**, Wu MT, Miller D, Wain JC, Kelsey KT, Wiencke JK, Christiani DC. The p53 codon 72 polymorphism and lung cancer risk. *Cancer Epidemiol Biomarkers Prev* 2000;**9**:1037-42.
22 **Prosser J**, Elder PA, Condie A, MacFadyen I, Steel CM, Evans HJ. Mutations in p53 do not account for heritable breast cancer: a study in five affected families. *Br J Cancer* 1991;**63**:181-4.
23 **Borresen AL**, Andersen TI, Garber J, Barbier-Piraux N, Thorlacius S, Eyfjord J, Ottestad L, Smith-Sorensen B, Hovig E, Malkin D, *et al.* Screening for germ line TP53 mutations in breast cancer patients. *Cancer Res* 1992;**52**:3234-6.
24 **Jolly KW**, Malkin D, Douglass EC, Brown TF, Sinclair AE, Look AT. Splice-site mutation of the p53 gene in a family with hereditary breast-ovarian cancer. *Oncogene* 1994;**9**:97-102.
25 **Sun XF**, Johannsson O, Hakansson S, Sellberg G, Nordenskjold B, Olsson H, Borg A. A novel p53 germline alteration identified in a late onset breast cancer kindred. *Oncogene* 1996;**13**:407-11.
26 **Warren W**, Eeles RA, Ponder BA, Easton DF, Averill D, Ponder MA, Anderson K, Evans AM, DeMars R, Love R, *et al.* No evidence for germline mutations in exons 5-9 of the p53 gene in 25 breast cancer families. *Oncogene* 1992;**7**:1043-6.
27 **Lehman TA**, Haffty BG, Carbone CJ, Bishop LR, Gumbs AA, Krishnan S, Shields PG, Modali R, Turner BC. Elevated frequency and functional activity of a specific germ-line p53 intron mutation in familial breast cancer. *Cancer Res* 2000;**60**:1062-9.
28 **Rapakko K**, Allinen M, Syrjakoski K, Vahteristo P, Huusko P, Vahakangas K, Eerola H, Kainu T, Kallioniemi OP, Nevanlinna H, Winqvist R. Germline TP53 alterations in Finnish breast cancer families are rare and occur at conserved mutation-prone sites. *Br J Cancer* 2001;**84**:116-19.
29 **Lidereau R**, Soussi T. Absence of p53 germ-line mutations in bilateral breast cancer patients. *Hum Genet* 1992;**89**:250-2.
30 **Shih HA**, Nathanson KL, Seal S, Collins N, Stratton MR, Rebbeck TR, Weber BL. BRCA1 and BRCA2 mutations in breast cancer families with multiple primary cancers. *Clin Cancer Res* 2000;**6**:4259-64.
31 **den Dunnen JT**, Antonarakis SE. Mutation nomenclature extensions and suggestions to describe complex mutations: a discussion. *Hum Mutat* 2000;**15**:7-12.
32 **Mehta CR**, Patel NR, Gray R. Computing an exact confidence interval for the common odds ratio in several 2x2 contingency tables. *J Am Stat Assoc* 1985;**80**:969-73.
33 **Toguchida J**, Yamaguchi T, Dayton SH, Beauchamp RL, Herrera GE, Ishizaki K, Yamamuro T, Meyers PA, Little JB, Sasaki MS, *et al.* Prevalence and spectrum of germline mutations of the p53 gene among patients with sarcoma. *N Engl J Med* 1992;**326**:1301-8.
34 **Brose MS**, Rebbeck TR, Calzone KA, Stopher JE, Nathanson KL, Weber BL. Cancer risk estimates for BRCA1 mutation carriers identified in a risk evaluation program. *J Natl Cancer Inst* 2002;**94**:1365-72.
35 **Storey A**, Thomas M, Kalita A, Harwood C, Gardiol D, Mantovani F, Breuer J, Leigh IM, Matlashewski G, Banks L. Role of a p53 polymorphism in the development of human papillomavirus-associated cancer. *Nature* 1998;**393**:229-34.
36 **Makni H**, Franco EL, Kaiano J, Villa LL, Labrecque S, Dudley R, Storey A, Matlashewski G. P53 polymorphism in codon 72 and risk of human papillomavirus-induced cervical cancer: effect of inter-laboratory variation. *Int J Cancer* 2000;**87**:528-33.
37 **Agorastos T**, Lambropoulos AF, Constantinidis TC, Kotsis A, Bontis JN. p53 codon 72 polymorphism and risk of intra-epithelial and invasive cervical neoplasia in Greek women. *Eur J Cancer Prev* 2000;**9**:113-18.
38 **Dokianakis DN**, Spandidos DA. P53 codon 72 polymorphism as a risk factor in the development of HPV-associated cervical cancer. *Mol Cell Biol Res Commun* 2000;**3**:111-14.
39 **Tenti P**, Vesentini N, Rondo Spaudo M, Zappatore R, Migliora P, Carnevali L, Ranzani GN. p53 codon 72 polymorphism does not affect the risk of cervical cancer in patients from northern Italy. *Cancer Epidemiol Biomarkers Prev* 2000;**9**:435-8.
40 **Tachezy R**, Mikyskova I, Salakova M, Van Ranst M. Correlation between human papillomavirus-associated cervical cancer and p53 codon 72 arginine/proline polymorphism. *Hum Genet* 1999;**105**:564-6.
41 **Yamashita T**, Yaginuma Y, Saitoh Y, Kawai K, Kurakane T, Hayashi H, Ishikawa M. Codon 72 polymorphism of p53 as a risk factor for patients with human papillomavirus-associated squamous intraepithelial lesions and invasive cancer of the uterine cervix. *Carcinogenesis* 1999;**20**:1733-6.
42 **Klaes R**, Ridder R, Schaefer U, Benner A, von Knebel Doeberitz M. No evidence of p53 allele-specific predisposition in human papillomavirus-associated cervical cancer. *J Mol Med* 1999;**77**:299-302.
43 **Rosenthal AN**, Ryan A, Al-Jehani RM, Storey A, Harwood CA, Jacobs IJ. p53 codon 72 polymorphism and risk of cervical cancer in UK. *Lancet* 1998;**352**:871-2.
44 **Ouchi T**, Monteiro AN, August A, Aaronson SA, Hanafusa H. BRCA1 regulates p53-dependent gene expression. *Proc Natl Acad Sci USA* 1998;**95**:2302-6.

45 **Zhang H**, Somasundaram K, Peng Y, Tian H, Bi D, Weber BL, El-Deiry WS. BRCA1 physically associates with p53 and stimulates its transcriptional activity. *Oncogene* 1998;**16**:1713-21.
46 **Marmorstein LY**, Kinev AV, Chan GK, Bochar DA, Beniya H, Epstein JA, Yen TJ, Shiekhattar R. A human BRCA2 complex containing a structural DNA binding component influences cell cycle progression. *Cell* 2001;**104**:247-57.
47 **Chai YL**, Cui J, Shao N, Shyam E, Reddy P, Rao VN. The second BRCT domain of BRCA1 proteins interacts with p53 and stimulates transcription from the p21WAF1/CIP1 promoter. *Oncogene* 1999;**18**:263-8.
48 **Harkin DP**, Bean JM, Miklos D, Song YH, Truong VB, Englert C, Christians FC, Ellisen LW, Maheswaran S, Oliner JD, Haber DA. Induction of GADD45 and JNK/SAPK-dependent apoptosis following inducible expression of BRCA1. *Cell* 1999;**97**:575-86.
49 **Crook T**, Crossland S, Crompton MR, Osin P, Gusterson BA. p53 mutations in BRCA1-associated familial breast cancer. *Lancet* 1997;**350**:638-9.
50 **Schuyer M**, Berns EM. Is TP53 dysfunction required for BRCA1-associated carcinogenesis? *Mol Cell Endocrinol* 1999;**155**:143-52.
51 **Bertwistle D**, Ashworth A. Functions of the BRCA1 and BRCA2 genes. *Curr Opin Genet Dev* 1998;**8**:14-20.
52 **Xu X**, Qiao W, Linke S, Cao L, Li W, Furth P, Harris C, Deng C. Genetic interactions between tumor suppressors Brca1 and p53 in apoptosis, cell cycle and tumorigenesis. *Nat Genet* 2001;**28**:266-71.
53 **Evans SC**, Mims B, McMasters KM, Foster CJ, de Andrade M, Amos CI, Strong LC, Lozano G. Exclusion of a p53 germline mutation in a classic Li-Fraumeni syndrome family. *Hum Genet* 1998;**102**:681–6.

Letter to the Editor

Am. J. Hum. Genet. 72:000, 2003

Resolving *ATM* Haplotypes in Whites

To the Editor:

In two recent studies, Bonnen et al. and Thorstenson et al. demonstrated extensive linkage disequilibrium distributed along *ATM* (GenBank accession number U82828) using SNPs (Bonnen et al. 2000; Thorstenson et al. 2001). In whites (Europeans), no recombination was observed along *ATM*. However, there are some discrepancies between the two articles, in that Thorstenson et al. found three haplotypes (H2, H3, and H4) in whites, whereas Bonnen et al. found five major haplotypes with frequencies >5% (2, 3, 15, 17, and 22) and two minor haplotypes in whites. Thorstenson et al. suggest that haplotype 2 and H4 may be equivalent, as may be 17 and H2 and 15 and H3. However, haplotype 22, which accounts for 35% of the population as determined by Bonnen et al., was not accounted for in the suggested equivalency. The two studies had only one overlapping SNP used to determine haplotypes, which may contribute to the discrepancy. Because we are interested in haplotyping *ATM* for association studies with breast cancer, we compared the haplotypes from the two studies in 159 individuals from 83 unrelated families with deleterious *BRCA1* mutations in order to determine which haplotypes are equivalent. Of the probands from the 83 families, all of whom carried *BRCA1* mutations, 72 were affected with cancer and 11 were not affected with cancer. In addition, we sought to determine the association of the three previously studied nonconservative coding region SNPs (S49C, D1853N, and P1054R) with each haplotype. As delineated by Thorstenson et al., there is a total of 12 nonconservative coding region SNPs in all populations, of which four appear in whites (S49C, F868L, D1853N, and P1054R). We were particularly interested in examining the association of the SNPs with the haplotypes, as Thorstenson et al. found that D1853N defined a single haplotype (H3), unlike Bonnen et al., who describe haplotype 15 (the suggested equivalent of H3) independently.

Bonnen typed 295 individuals from four ethnic groups (71 African Americans, 39 Asian Americans, 77 white European Americans, and 73 Hispanic Americans) for 14 SNPs that spanned 142 kb across *ATM*. Using the 14 SNPs, they predicted a total 22 of *ATM* haplotypes, using EMHAPFRE, with five predominant haplotypes having a frequency ≥5%. The major haplotypes identified in white European Americans were 2 (29.2%), 3 (6.5%), 15 (17.5%), 17 (10%), and 22 (35.1%), as shown in table 1. In addition, they examined the association between three nonconservative coding region SNPs (S49C, D1853N, and P1054R) and the haplotypes they determined. Each nonconservative coding region SNP showed a significant association with a specific haplotype of *ATM*, as defined in their study (table 2). SNP1 (S49C) showed an association with haplotype 2, SNP2 (D1853N) with haplotype 15, and SNP3 (P1054R) with haplotype 17.

Thorstenson et al. typed 93 individuals from seven major human populations (18 from Africa, 9 from the Middle East, 12 from the Indian peninsula, 20 from Asia, 16 from Europe, 8 from Oceania, and 10 American Indians) for 17 SNPs (only one common to the 14 SNPs in the work of Bonnen et al.) spanning 146 kb across *ATM*. Ten of the 17 SNPs were found to be in complete linkage disequilibrium and were used to construct the *ATM* haplotypes. Seven haplotypes (H1–H7) were inferred using a maximum parsimony approach. In the European population, three major haplotypes were identified: H2 (40%), H3 (12.5%), and H4 (47%) (table 1).

Table 1

Frequency of Haplotypes from Bonnen et al., Thorstenson et al., and Current Study Determined in Probands from Families and All Family Members

Study and Haplotypes	Frequency (%)		
		Current Study	
	Published	Probands	All Individuals
Bonnen et al.:			
2	29	33	39
3	6.5	4	3
15	17.5	16	14
17	10	12	14
22	35	35	30
Thorstenson et al.:			
H2	40	47	43
H3	12.5	17	14
H4	47	37	43

Thorstenson et al. also examined the association between these haplotypes and the same amino acid variant SNPs as Bonnen et al. S49C showed an association with H4, P1054R showed an association with H2, and D1853N defined H3.

For our comparison, we constructed new *ATM* haplotypes by genotyping 159 individuals (318 alleles) from 83 families with deleterious *BRCA1* mutations; 150 were white (non-Hispanic), and 9 were African American. These individuals are representative of our breast cancer study population. Our aim was to reconstruct *ATM* haplotypes as closely equivalent as possible to those of the other two studies, using the minimum number of SNPs from each paper that defined each haplotype. Thus, the following SNPs were genotyped: 10182, IVS46-257, IVS55+186, and IVS62-694 from the Bonnen study and IVS17-56 and D1853N from the Thorstenson study. The SNPs selected for this study allowed definition of all the major haplotypes in whites with haplotype frequencies >5%. Haplotypes 6 and 21, seen in table 2 of Bonnen et al., have haplotype frequencies <5% and, therefore, were not included in the study. Of the 159 samples typed using the SNPshot protocol on an ABI Prism 3100, all but two samples (1%) were consistent with the haplotype equivalencies shown in figure 1.

On the basis of our findings (shown in fig. 1), Bonnen's haplotype 22 and haplotype 17 are encompassed by Thorstenson's H2, and haplotypes 2 and 3 are encompassed by H4. Haplotype 15 is equivalent to H3. Our haplotype frequencies are consistent with those of Bonnen and Thorstenson in white individuals, as shown in table 1. For the two samples that did not fit into the equivalencies suggested in figure 1, one of the two samples contained haplotype 11, which was shown by Bonnen et al. to have a 1.3% frequency in the Asian population. It was seen in an individual homozygous for H4 and appears to be derived from haplotype 2. In the other sample, the haplotypes were not resolvable despite repeated genotyping.

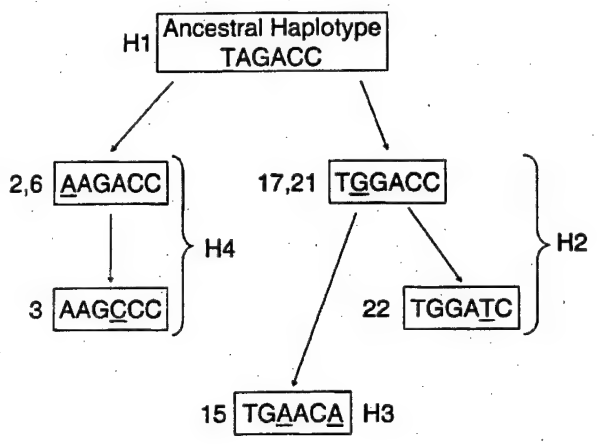

Figure 1 Phylogenetic relationship and equivalencies among major haplotypes in the work of Bonnen et al. and Thorstenson et al. Base pair changes defining the haplotypes are 10182T→A, IVS17-56G→A, 5557G→A (D1853N), IVS46-257A→C, IVS55+186C→T, and IVS62-694C→A.

As in the study by Bonnen et al, we report the percentage of the most frequent haplotype of the total number of alleles that are in the individuals with three nonconserved coding region SNPs (table 2). Similar to the study of Bonnen et al., haplotype 2 is the most frequently occurring allele in the individuals with S49C (cSNP1; 50%) but does not differ significantly from the percentage of haplotype 2 (of the total alleles) in the remaining individuals (38%; $P = .6$). Haplotype 17 is the most frequent allele in the individuals with P1054R (cSNP3; 46%), significantly more than in the alleles of the individuals without P1054R (9%; $P = .003$). Neither S49C nor P1054R is found exclusively on the haplotypes they are most frequently associated with, 2 and 17, respectively. For both S49C and P1054R, if the individual did not carry the most frequent haplotype (i.e., 2 and 17, respectively), he or she carried the haplotype

Table 2

Most Frequent Haplotypes within Individuals Carrying Nonconservative Coding Region SNPs

	Bonnen et al.			Thorstenson et al.	
Nonconservative Coding Region SNPs	Associated Haplotype	Frequency of Associated Haplotype (%)		Associated Haplotype[b]	Frequency in Current Study of Associated Haplotype[b] (%)
		Current Study[a]	Published[a]		
S49C	2	50	64	H4	56
D1853N	15	66	57	H3	66
P1054R	17	46	52	H2	57

[a] Compares the percentage of the alleles with the most frequent haplotype in the group of individuals with the SNP, as published by Bonnen et al. and in the current study.

[b] Shows the percentage of total alleles contributed by the equivalent haplotypes (from Thorstenson et al.) in all the individuals with the SNP.

derived from the most common haplotype (i.e., 3 and 22, respectively).

Unlike Bonnen et al., we did not find haplotype 15 in individuals without the D1853N SNP (cSNP2), and the frequency of haplotype 15 in the individuals with D1853N is entirely reflective of the rate of heterozygotes and homozygotes for D1853N. Our results are consistent with those of Thorstenson et al., who found the 1853N SNP defining a specific haplotype (H3). However, in general, our results are similar to those found by Bonnen et al. in the white population. In light of the interest in completing haplotype maps of the genome, this study illustrated two points that need to be taken into consideration in haplotype-association studies. First, haplotype association studies might miss functional SNPs similar to S49C, since haplotype 2 is no more frequent in carriers of S49C than noncarriers of S49C. Secondly, some nonconservative coding region SNPs, although associated with certain haplotypes, are not always seen in the context of the same haplotype, as seen with S49C and P1054R, whereas others are completely associated, as seen with D1853N. Our observation illustrates the importance of constructing phylogenetic trees to understand how haploypes might be grouped together for association studies. Thus, association studies using haplotype maps need to be constructed carefully with thought to the potential pitfalls demonstrated by this study.

Acknowledgments

We acknowledge the helpful comments of the two anonymous reviewers of this manuscript. This study was supported by the National Institutes of Health Grant K08 CA84030 (to K.L.N.).

RICHARD LETRERO,[1] BARBARA L. WEBER,[1,2] AND KATHERINE L. NATHANSON[1]

[1]*Department of Medicine and*
[2]*Abramson Family Cancer Research Institute*
University of Pennsylvania School of Medicine
Philadelphia

Electronic-Database Information

Accession number and URL for data presented herein are as follows:

GenBank, http://www.ncbi.nlm.nih.gov/Genbank/ (for genomic sequences of *ATM* [accession number U82828])

References

Bonnen PE, Story MD, Ashorn CL, Buchholz TA, Weil MM, Nelson DL (2000) Haplotypes at *ATM* identify coding-sequence variation and indicate a region of extensive linkage disequilibrium. Am J Hum Genet 67:1437–1451

Thorstenson YR, Shen P, Tusher VG, Wayne TL, Davis RW, Chu G, Oefner PJ (2001) Global analysis of *ATM* polymorphism reveals significant functional constraint. Am J Hum Genet 69:396–412

Address for correspondence and reprints: : Dr. Katherine L. Nathanson, 513 BRB 2/3, 421 Curie Boulevard, University of Pennsylvania School of Medicine, Philadelphia, PA 19104. E-mail: knathans@mail.med.upenn.edu

0002-9297/2003/7204-00XX$15.00

Original Paper

Human Heredity

Hum Hered 2003;55:171–178
DOI: 10.1159/000073201

Received: December 10, 2003
Accepted after revision: May 27, 2003

Population Frequencies of Single Nucleotide Polymorphisms (SNPs) in Immuno-Modulatory Genes

A.-M. Martin[a,b,c] G. Athanasiadis[b,c] J.D. Greshock[b,c] J. Fisher[b] M.P. Lux[e]
K. Calzone[c] T.R. Rebbeck[d,e] B.L. Weber[b,c]

[a]Laboratory of Molecular Pathology, Department of Pathology, Pennsylvania Hospital, [b]Department of Medicine, University of Pennsylvania, [c]Abramson Family Cancer Research Institute, University of Pennsylvania Cancer Center, [d]Department of Biostatistics and Epidemiology, and [e]Center for Clinical Epidemiology and Biostatistics, University of Pennsylvania, Philadelphia, Pa., USA

Key Words
Cytokines · SNPs · Allele frequencies · African American · Caucasian

Abstract
Inherited polymorphisms in immuno-modulatory genes may contribute to variations in immune function and genetic susceptibility for complex diseases, including cancer. We report results from a comprehensive study to discover novel single nucleotide polymorphisms (SNPs) and to estimate allelic frequency for both novel and known coding and regulatory region SNPs in genes encoding proteins that have been implicated in the immune response to tumors. We identified 12 novel nucleotide substitution variants and one deletion variant in 17 genes analyzed (*TGFβR, β2M, IFNγ, TNFα, TNFαR, LTα, IL-6, IL-12, IL-2, IL-1α, IL-1β, IL-1RN, IL-10, CTLA4, CD40L, Fas* and *FasL*). We determined the frequency of these novel polymorphisms, as well as 17 previously identified polymorphisms, in a control sample of 158 individuals, approximately half of which were Caucasian (n = 74) and half of which were African American (n = 84). Significant differences in allele frequencies were observed between the two racial groups for 13/17 genes tested. These allelic variations maybe associated with alterations in immune function and thus susceptibility to a number of complex disease states such as cancer.

Introduction

The immune system is a complex network of cells that has evolved to protect humans against infectious agents and tumor growth. In the case of cancer, the major antitumor effect is cell-mediated with the involvement of T lymphocytes, as well as natural killer (NK) cells. T lymphocytes are activated by recognizing peptides or antigens on the surface of target cells in the context of the major histocompatibility complex. Following activation of antigen-specific T cells; a cascade of events takes place leading to proliferation of the T cell itself, as well as recruitment of other immune cells and secretion of cytokines. The central role of cytokines as mediators of the immune response, as well as their involvement in various immuno-

KARGER
Fax + 41 61 306 12 34
E-Mail karger@karger.ch
www.karger.com

Accessible online at:
www.karger.com/hhe

Barbara Weber, MD
Room 514, BRB II/III, 421 Curie Boulevard
Philadelphia, PA 19104 (USA)
Tel. +1 215 898 0247, Fax +1 215 573 2486, E-Mail weberb@mail.med.upenn.edu

logical functions, is of interest to many investigators. The balance between pro- and anti-inflammatory cytokines is critical for the immune system to function adequately and can greatly influence the outcome of the immune response to protect against disease development. Thus, specific immuno-modulatory gene polymorphisms associated with gene regulation and protein expression may influence clinical outcome of a number of disease states, such as malignancy [1, 2], infectious diseases [3, 4], autoimmunity [5], transplant tolerance [6], asthma and allergy [60] and graft-versus-host disease [7].

Familial studies of cancer have identified a number of cancer susceptibility genes, including high penetrance genes (*BRCA1*, [8]) and low penetrance genes (CHEK2 [9], androgen receptor and N-acetytransferase 1(reviewed in [10]). Low penetrance genes may modify high penetrance genes [reviewed in 11, 12], for example specific variant forms of *MC1R* modify *CDKN2A* penetrance and the development of melanoma [13]. It is postulated that low penetrance cancer susceptibility alleles are the result of genetic polymorphisms (the most frequent of which are single nucleotide polymorphisms (SNPs)), which contribute to variation in gene expression or function and thus alter risk for disease [11, 14]. One example of the functional effect of genetic variation may be the documented inter-individual differences in immune response capacity [15, 16]. This hypothesis is supported by the association of several known SNPs in immuno-modulatory genes and altered immune capacity [reviewed in 17, 18]. Thus, given that tumor development may in part depend on escaping immune surveillance, these SNPs may constitute risk factors for cancer development.

There is evidence to show that the random distribution of allele frequencies throughout the human genome follows diverse ethnic and/or racial trends [19, 20]. The frequency of sequence variations can differ by race and ethnicity and this variation may be associated with a difference in risk for disease between these groups [11]. For example, documented differences in allele frequencies between African Americans and Caucasians for genes involved in DNA repair [21] and hormone metabolism [22] have been proposed to contribute to differences in lung cancer [23] breast cancer [21, 24] and prostate cancer risk [25]. Inter-racial studies of immune function suggest that there are differences between African Americans and Caucasians in leukocyte subsets [26] as well as in the expression of co-stimulatory molecules on the surface of lymphocytes. Furthermore, up to 30% of healthy Caucasians have a constitutively low natural killer cell count [27], likely a result of polymorphic genetic variants.

As noted above, several studies have reported associations between immune response gene variants and disease susceptibility; however, a comprehensive survey of candidates for antitumor immune response genes and an analysis of the underlying frequency and distribution of these gene variants across populations have not been performed. We analyzed a panel of 17 genes including, cytokines such as interleukins *(IL-1, IL-2, IL-12, IL-6, IL-10)*, interferon gamma *(IFNγ)*, tumor necrosis factor alpha *(TNFα)*, lymphotoxin *(LTα)*, the co-stimulatory molecules *(CTLA4, CD40L)* and apoptotic factors *(Fas, Fas-ligand)*. Our aim was to (i) identify novel sequence variants in this panel and (ii) to determine allele frequencies in a control sample of Caucasians and African Americans for both novel and previously identified polymorphisms in these genes. Our rationale for choosing this panel was based on known function and association with anti-tumor immunity, but the same genes are relevant to the study of autoimmune diseases, infectious diseases and transplantation immunology as well. All sequence variants identified were located in coding and regulatory regions of the immuno-modulatory genes.

Materials and Methods

Population Samples

One hundred and fifty individuals were analyzed in this study. Eighty-four African Americans and 36 Caucasians were ascertained from the referral regions for the University of Pennsylvania Health System (UPHS) using flyers, radio announcements and newspaper advertisements. An additional 38 Caucasians (males = 25 and females = 13) were ascertained as spouses of individuals seen for breast cancer risk assessment at the Cancer Risk Evaluation Program of the University of Pennsylvania (1994–1998) or at the University of Michigan (1993–1994) for clinical research studies. A peripheral blood sample was collected from each individual for DNA preparation after obtaining informed consent. Individuals were included for study participation if they had no prior personal or family history of cancer of any kind. In addition, all participants were over the age of 18 years and of Caucasian or African American ancestry. The study was reviewed and approved by the Institutional Review Boards of both participating institutions.

PCR Amplification

DNA was extracted from peripheral blood mononuclear cells and stored in TE at 4°C. The coding and regulatory regions of the genes tested were PCR amplified (primer sequences and conditions can be found at URL: http://linkage.rockefeller.edu/hh/martinprimers.html). PCR amplification was performed in a final volume of 20 µl containing 80 ng of DNA, 1.5 m*M* of $MgCl_2$, 10 m*M* Tris-HCl (pH 8.3), 50 m*M* KCl, 0.2 m*M* each of dCTP, dATP, dTTP, dGTP (Amersham Pharmacia Biotech), each primer at 1.0 µ*M* and 1.0 unit of Taq polymerase (Boehringer Mannheim). Amplifications were performed in a 9700 Perkin Elmer/Cetus Thermocycler.

Table 1. Novel polymorphisms in immune response genes

Gene	Gene position	Nucleotide substitution		Amino acid substitution	
		nucleotide	change	codon	variant
TNFα	5′ UTR	−106	del G	–	–
TNFα	5′ UTR	−28	T→C	–	–
TNFαR	Exon 3	+7979	C→T	75	P75L
IL-1α	Intron 1	−35	G→A	–	–
IL-1α	Exon 4	+2121	C→T	92	–
IL-12p35	Promoter	−1250	T→A	–	–
IL-12p35	Promoter	−666	T→G	–	–
IL-12p40	Promoter	−5230	A→G	–	–
IL-12p40	Promoter	−5251	C→T	–	–
IL-12p40	Promoter	−3882	A→G	–	–
IL-12p40	Promoter	−5310	T→A	–	–
CTLA4	Exon 2		G→A	90	M90I

Nucleotide position based on genomic DNA sequence.
All nucleotide positions calculated from translation start site = +1 [63].

PCR products were analyzed by one of three mutation detection methods: conformation sensitive gel electrophoresis (CSGE) [28] and subsequent confirmation of variants by direct sequencing; second, by restriction fragment length polymorphism (RFLP) analysis and subsequent confirmation of variants by direct sequencing and third, by direct sequencing after PCR amplification of the regulatory and coding regions of the gene.

CSGE

PCR products were denatured at 98°C for 5 min and then re-annealed at 68°C for 30 min to allow heteroduplex formation. Gels consisted of 0.5× Tris-Taurate EDTA (TTE) buffer (44.4 m*M* Tris/14.5 m*M* Taurine (USB)/1.0 m*M* EDTA, pH 9.0, filter), 10% polyacrylamide with 99:1 ratio of acrylamide to 1,4-bis(acryloyl) piperazine (BAP; Fluka), 15% formamide, 10% ethylene glycol, 0.1% ammonium persulphate and 0.69% N,N,N′,N′,-tetramethylethylenediamine (TEMED). Gels were run overnight at 10–25 W, depending upon the length of the PCR product. Gels were stained with ethidium bromide for 10–15 min and visualized by UV light.

RFLP Analysis

For some assays, 10 μl of the PCR product was digested with 2 units of the appropriate restriction enzyme using the manufacturer's recommended protocol. Visualization of the PCR products was accomplished using 1–3% agarose gel electrophoresis stained with ethidium bromide.

Sequence Analysis

Sequence variants were re-amplified from source DNA as described above. The PCR products were purified by QIAquick PCR purification kit (Qiagen) or by ExoSapTM following the manufacturer's guidelines. PCR products were subsequently sequenced using the ABI prism 377 and the DyeTerminatorTM kit (Perkin Elmer) according to manufacturer's guidelines. All unique or rare variants were sequenced directly.

Statistical Analysis

The JMP and SAS statistical analysis packages were used for all analyses. Allele frequencies were estimated by allele counting methods and differences in allele frequencies between the two populations were determined using a 2 × 3 contingency table and Fishers Exact Method using the JMP statistical package. Hardy-Weinberg estimates were performed using the SAS statistical package.

Results

In order to identify novel polymorphisms we screened the coding region of 14 genes (see tables 1–3) in 158 control samples using CSGE. In addition, we screened the promoters and/or 5′ UTR regions for 10/14 genes (*TNFα, IL-1α, IL-2, CTLA4, IFNγ, IL-12, Fas, FasL, IL-12p35, IL-12p40,* and *CD40L*). Finally, in a further 3 genes we determined only the 252G→A, −236G→C (−174G→C) polymorphism and the −854C→T (−819C→T) polymorphism in *LTα, IL-6* and *IL-10* respectively. We did not detect any known or novel polymorphisms in *TGFβRI, CD40L, β2M, IFNγ* and *Fas-L.*

Novel Polymorphisms

Table 1 lists all novel polymorphisms identified in this study. Eleven novel nucleotide substitution variants and one single nucleotide deletion variant were identified. We found 8/12 (67%) novel nucleotide substitutions in regulatory regions, 1/12 substitutions (8%) in intronic regions and the remaining 3/12 (25%) in coding regions of the genes. Only two novel coding region nucleotide substitutions resulted in an alteration in the amino acid sequence

Table 2. Previously identified polymorphisms in immune response genes

Gene	Gene position	Nucleotide substitution		Amino acid substitution	
		nucleotide	change	codon	variant
TNFα	Promoter	−418 (−238)	G→A	–	–
TNFα	Promoter	−488 (−308)	G→A	–	–
TNFα	Promoter	−1021 (−850)	C→T	–	–
TNFα	Promoter	−1027 (−856)	C→A	–	–
TNFαR	Exon 1	+36	G→A	12	–
IL-1β	Exon 5	+3406 (+3953)	T→C	105	–
IL-1RN	Exon 2	+2001 (+8006)	T→C	57	–
IL-1RN	Exon 4	+4095 (+11100)	T→C	130	–
IL-2	Exon 1	+112 (+742)	T→G	38	–
IL-2	Intron 2	+360	C→A	–	–
IL-6	Promoter	−236 (−174)	G→C	–	–
IL-10	Promoter	−854 (−819)	C→T	–	–
CTLA4	Promoter	−318	C→T	–	–
CTLA4	Exon 1	49	A→G	17	T17A
LTα	Intron 1	252	G→A	–	–
Fas	Exon 3	174	A→G	58	–
Fas	Intron 5	439	G→C	–	–

Nucleotide position based on genomic DNA sequence.
All nucleotide positions calculated from translation start site = +1 [63].
Numbers in parentheses correspond to previously published nucleotide positions.

(*TNFαR*, P75L and *CTLA4*, M90I; see table 1), the remaining novel coding region substitution was silent (*IL-1α* 2121C→T; see table 1). None of the substitutions destroyed a splice site or generated a cryptic splice site.

Previously Identified Polymorphisms

Table 2 lists all 17 previously identified polymorphisms detected in our study. Seven (41%) nucleotide substitutions were detected in the promoter region, 3/17 (18%) were detected in the intronic regions and 7/17 (41%) were found in the exonic regions of the genes. Only one of the exonic polymorphisms altered an amino acid sequence (*CTLA4* T17A) and the remaining six were silent (see table 2).

Coding Region Polymorphisms

Three of ten exonic substitutions were novel (*TNFαR*, *CTLA4* and *IL-Iα*); the remaining seven exonic substitutions were previously reported [29–33] (see tables 1, 2). Three of ten exonic SNPs altered the amino acid sequence; the remaining seven were silent. One novel missense substitution was identified in the *TNFαR* gene (7979C→T), resulting in the substitution of proline to leucine in codon 75. The two remaining missense substitutions were both found in the *CTLA4* gene. The previously documented *CTLA4*, 49A→G nucleotide substitution 37 results in a substitution from threonine to alanine in codon 17 and this alteration changes the polarity of the amino acid. Finally, the novel *CTLA4*, 2841G→A missense substitution results in a substitution of methionine to isoleucine in codon 90.

Noncoding Region Polymorphisms

Two novel polymorphisms were detected 5′ of the translation start site in the *TNFα* gene, one of which was a deletion variant (−106delG) (table 2). Four previously identified nucleotide substitutions also were detected in the promoter of the *TNFα* gene [34, 35]. All six SNPs identified in the *IL-12* promoter were novel and the remaining three SNPs in the promoters of *CTLA4, IL-10* and *IL-6* had been previously identified [36, 37]. A total of four intronic SNPs were detected in this study; one novel non-coding SNP was identified in intron 1 of *IL-1α* and three previously identified SNPs in intron 1 of *LTα intron 2 of IL2i* and intron 5 of *Fas* respectively [33, 38, 39].

Polymorphism Frequencies

Variant alleles existed in frequencies ranging from 1 to 60% in this sample of 158 individuals (table 3). There was

Table 3. Frequency of immune response alleles in Caucasians and African Americans

Gene	Polymorphism[a]	Genotype frequency among Caucasians (n = 74)			Genotype frequency among African Americans (n = 84)			Difference between racial groups p value*
		WT/WT %	WT/V %	V/V %	WT/WT %	WT/V %	V/V %	
TNFα	−418G→A	91	4	5	90	5	5	–
	−488G→A	87	12	1	87	12	1	–
	−1021C→T	74	23	3	93	6	1	0.03
	−1027C→A	88	11	1	79	20	1	–
	−106delG	100	–	–	99	1	–	–
	−28T→C	100	–	–	96	4	–	0.0001
TNFαR	36G→A	81	19	–	100	–	–	0.0001
	7979C→T	100	–	–	85	15	1	0.0002
IL-1α	−35G→A	99	1	–	87	13	–	–
	2121C→T	99	1	–	88	7	5	0.03
IL-1β	3406C→T	60	35	5	79	19	2	0.033
IL-1RN	2001T→C	46	45	9	77	23	–	0.0003
	4095T→C	51	41	8	72	24	4	0.008
IL-2	112T→G	49	45	6	73	24	3	0.009
	360C→A	100	–	–	87	12	1	0.006
IL-6	−236G→C	40	46	14	84	16	–	0.0001
IL-10	−854C→T	59	34	7	42	49	9	–
IL-12p35	−1250T→A	100	–	–	99	1	–	–
	−666T→G	80	19	1	91	9	–	–
IL-12p40	−5230A→G	100	–	–	92	8	–	0.015
	−5251C→T	99	1	–	93	7	–	–
	−3882A→G	100	–	–	99	1	–	–
	−5310T→A	97	3	–	89	11	–	–
CTLA4	−318C→T	76	24	–	99	1	–	0.0001
	49G→A	45	39	16	32	45	23	–
	2841G→A	100	–	–	99	1	–	–
LTα	252G→A	41	54	5	23	56	21	0.0001
Fas	174A→G	87	10	3	91	5	4	–
	439G→C	63	34	3	84	6	–	

[a] Nomenclature = nucleotide position in relation to translation start site (den Dunnen and Antonarakis, 2001 [63].
* All p values were calculated by 2 × 3 contingency tables using JMP statistics package; only statistically significant p values are listed.

a statistically significant difference in the allele frequency for 14/29 polymorphisms between Caucasians and African Americans (table 3). Of note, 8/14 novel sequence variations identified were seen exclusively among the African American samples (see table 3). Only the *TNFαR*, 36G→A sequence variation was seen exclusively among the Caucasian samples. Finally, all polymorphisms where with sufficient numbers of variants for evaluation adhered to Hardy-Weinberg equilibrium as determined by χ^2 analysis.

Discussion

This study was undertaken as a comprehensive survey of genes involved in anti-tumor immunity to identify novel genetic polymorphisms and to define allele frequencies of novel and previously described polymorphisms in both Caucasians and African Americans in immuno-modulatory genes. Most cytokines and immuno-modulatory genes are regulated predominantly at the transcriptional level [40, 41]. Transcription of these genes involves the cooperative interaction of cell-specific transcription factors bound to regulatory elements in the promoters. Thus, SNPs in the regulatory regions of the genes could subtly alter the transcriptional regulation of the genes. There are precedents for this model; as a number of the SNPs evaluated in this study not only affect gene expression but also are associated with disease characteristics (http://bris.ac.uk/pathandmicro/services/GAI/cytokine4.htm). For instance, the *IL-6* −236G→C (−174G→C) polymorphism is situated immediately 5′ to the multiple responsive element in the promoter [42]. Some studies indicate that the G-allele influences *IL-6* transcription [43] while others have concluded it does not [44]. However, the presence of the G-allele is associated with systemic onset juvenile chronic arthritis [43].

In addition the *TNFα* −488G→A (−308G→A) polymorphism lies in a consensus sequence for an AP2 binding site, and alters the ability of AP2 to bind to this site [45]. Conflicting results exist for the effect of this polymorphism and gene expression [46–50], nonetheless, homozygosity for the A-allele of this SNP is associated with a seven-fold increased risk of death from cerebral malaria [51], and the G-allele is associated with a 3-fold increased risk for developing chronic lymphocytic leukaemia [2]. In order to determine the function of the novel SNPs identified in our study, it will be necessary to conduct similar functional assays. In addition, large case-control association studies will provide insight into the role that these polymorphic genes play in disease risk.

Interestingly, 7/12 novel SNPs identified in this study were seen exclusively in the African American population and only one novel sequence variation was found exclusively in the Caucasian population. There are several possible explanations for these data including a sample size too small for rare polymorphisms comparison between populations or true specificity to the African American population. It is known that SNP allele frequencies vary considerably across human ethnic groups and populations [reviewed in 11]. The frequency of alleles and genotypes of the sequence variations in our two populations differed significantly for a number of genes tested. The difference in allele frequency was statistically significant for 14 of the sequence variations identified. These differences must be taken into account when performing studies of disease-risk and generalizing across populations.

We did not detect polymorphisms in five of the genes tested (*β2M, TGFβR11, CD40L, FasL* and *IFNγ*). Furthermore, a number of previously documented SNPs in *IFNγ, TNFα* and *Fas* were not detected using conventional CSGE, possibly due to low allele frequencies and lack of representation in this sample. CGSE is approximately 95% sensitive as compared to direct sequencing, thus it is less likely that this finding is due to technical insensitivity. Still, there have been recent improvements in SNP detection, including capillary electrophoresis (CSCE) detection [52] and high-performance liquid chromatography (HPLC) [reviewed in 53] that may increase the sensitivity of the assay [54, 55]. Most recently, CSCE was used to detect *RAS* and *BRAF* mutations in melanoma and lung carcinoma [61, 62], validating this technique as an important screening tool.

The balance between pro-inflammatory and anti-inflammatory T cells influences the pathology of malignant diseases. Inter-individual differences in immuno-modulatory gene profiles appear to be due, at least in part to allelic polymorphisms within the regulatory and coding regions of genes of the immune system. From the results we present here it is possible that African Americans have a genetically determined, quantitatively different immune response than Caucasians, which could contribute to adverse disease outcomes. African Americans are at a higher risk for rejecting allografts, have a poorer prognosis in breast cancer [56, 57] and other cancers [58] and an altered ability to recover from infectious diseases [59], also suggesting that African-American patients may form an immunologically higher risk group. Further studies to investigate this hypothesis will be required. These data will greatly facilitate efforts aimed at unraveling the complex traits governed by the involvement of the immune system and the differences between racial groups.

Acknowledgements

The authors express thanks to B.J. Baxter, P. Volpe and J. Harrow for their technical assistance. This work was supported by grants from the Susan G. Komen Breast Cancer Foundation (PDF99-003024; to A-M.M and BLW) and a grant from the NCI (CA82707 to BLW).

References

1 Demeter J, Messer G, Ramisch S, Mee JB, di Giovine FS, Schmid M, Herrmann F, Porzsolt F: Polymorphism within the second intron of the IL-1 receptor antagonist gene in patients with hematopoietic malignancies. Cytokines Mol Ther 1996;2:239–242.

2 Demeter J, Porzsolt F, Ramisch S, Schmidt D, Schmid M, Messer G: Polymorphism of the tumour necrosis factor-alpha and lymphotoxin-alpha genes in chronic lymphocytic leukaemia. Brit J Haem 1997;97:107–112.

3 Hurme M, Lahdenpohja N, Santtila S: Gene polymorphisms of interleukins 1 and 10 in infectious and autoimmune diseases. Ann Med 1998;30:469–473.

4 Knight JC, Kwiatkowski D: Inherited variability of tumor necrosis factor production and susceptibility to infectious disease. Proc Assoc Am Phys 1999;111:290–298.

5 Komata T, Tsuchiya N, Matsushita M, Hagiwara K, Tokunaga K: Association of tumor necrosis factor receptor 2 (TNFR2) polymorphism with susceptibility to systemic lupus erythematosus. Tiss Antigens 1999;53:527–533.

6 George S, Turner D, Reynard M, Navarrete C, Rizvi I, Fernando ON, Powis SH, Moorhead JF, Varghese Z: Significance of cytokine gene polymorphism in renal transplantation. Transplant Proc 2001;33:483–484.

7 Cavet J, Middleton PG, Segall M, Noreen H, Davies SM, Dickinson A: Recipient tumor necrosis factor-α and interleukin-10 gene polymorphisms associate with early mortality and acute graft-versus-host disease severity in HLA-matched sibling bone marrow transplants. Blood 1999;94:3941–3946.

8 Miki Y, Swensen J, Shattuck-Eidens D, Futreal PA, Harshman K, Tavtigian S, Liu Q, Cochran C, Bennett LM, Ding W: A strong candidate for the breast and ovarian cancer susceptibility gene BRCA1. Science 1994;266:66–71.

9 Meijers-Heijboer H, van den Ouweland A, Klijn J, Wasielewski M, de Snoo A, Oldenburg R, Hollestelle A, Houben M, Crepin E, van Veghel-Plandsoen M, Elstrodt F, van Duijn C, et al: Low-penetrance susceptibility to breast cancer due to CHEK2(*)1100delC in noncarriers of BRCA1 or BRCA2 mutations. Nat Genet 2002;31(1):55–9. Available from http://www.ncbi.nlm.nih.gov/htbin-post/Entrez/query?db=m&form=6&dopt=r&uid=11967536.

10 Martin AM, Weber BL: Genetic and hormonal risk factors in breast cancer. J Natl Cancer Inst 2000;92:1126–1135.

11 Shields PG, Harris CC: Cancer risk and low-penetrance susceptibility genes in gene-environment interactions. J Clin Oncol 2000;18: 2309–2315.

12 Weber BL, Nathanson KL: Low penetrance genes associated with increased risk for breast cancer. Eur J Cancer 2000;36:1193–1199.

13 Box NF, Duffy DL, Chen W, Stark M, Martin NG, Sturm RA, Hayward NK: MC1R genotype modifies risk of melanoma in families segregating CDKN2A mutations. Am J Hum Genet 2001;69:765–773.

14 Rebbeck T, Wang Y, Kantoff PW, Krithivas K, Neuhausen S, Godwin AK, Daly MB, Narod SA, Brunet J-S, Vesprini D, Garber JE, Lynch HT, Weber BL, Brown M: Modification of BRCA1- and BRCA2-associated breast cancer risk by AIB1 genotype and reproductive history. Cancer Res 2001;61(14):5420–5424.

15 Hutchings A, Purcell WM, Benfield MR: Increased costimulatory responses in African-American kidney allograft recipients. Transplantation 2001;71:692–695.

16 Hutchings A, Guay-Woodford L, Thomas JM, Young CJ, Purcell WM, Pravica V, Perrey C, Hutchinson IV, Benfield MR: Association of cytokine single nucleotide polymorphisms with B7 costimulatory molecules in kidney allograft recipients. Ped Transpl 2002;6:69–77.

17 Bidwell J, Keen L, Gallagher G, Kimberly R, Huizinga T, McDermott MF, Oksenberg J, McNicholl J, Pociot F, Hardt C, D'Alfonso S: Cytokine gene polymorphism in human disease: on-line databases, supplement 1. Genes Immun 2001;2:61–70.

18 Bidwell J, Keen L, Gallagher G, Kimberly R, Huizinga T, McDermott MF, Oksenberg J, McNicholl J, Pociot F, Hardt C, D'Alfonso S: Cytokine gene polymorphism in human disease: on-line databases. Genes Immun 1999;1: 3–19.

19 Garte S: The role of ethnicity in cancer susceptibility gene polymorphisms: the example of CYP1A1. Carcinogenesis 1998;19:1329–1332.

20 Weber W: Populations and genetic polymorphisms. Molecular Diagnosis 1999;4:299–307.

21 Duell E, Millikan R, Pittman G, Winkel S, Lunn R, Tse C, Eaton A, Mohrenweiser H, Newman B, Bell D: Polymorphisms in the DNA repair gene XRCC1 and breast cancer. Cancer Epidemiol Biomarkers Prev 2001;10: 217–222.

22 Kuehl P, Zhang J, Lin Y, Lamba J, Assem M, Schuetz J, Watkins P, Daly A, Wrighton S, Hall S, Maurel P, Relling M, et al: Sequence diversity in CYP3A promoters and characterization of the genetic basis of polymorphic CYP3A5 expression. Nat Genet 2001;27:383–391.

23 Wu X, Zhao H, Amos CI, Shete S, Makan N, Hong WK, Kadlubar FF, Spitz MR: p53 genotypes and haplotypes associated with lung cancer susceptibility and ethnicity. J Nat Cancer Inst 2002;94:681–690.

24 Bailey LR, Roodi N, Verrier CS, Yee CJ, Dupont WD, Parl FF: Breast cancer and CYPIA1, GSTM1, and GSTT1 polymorphisms: evidence of a lack of association in Caucasians and African Americans. Cancer Res 1998;58: 65–70.

25 Tang YM, Green BL, Chen GF, Thompson PA, Lang NP, Shinde A, Lin DX, Tan W, Lyn-Cook BD, Hammons GJ, Kadlubar FF: Human CYP1B1 Leu432Val gene polymorphism: ethnic distribution in African-Americans, Caucasians and Chinese; oestradiol hydroxylase activity; and distribution in prostate cancer cases and controls. Pharmacogenetics 2000;10:761–766.

26 Hutchings A, Purcell WM, Benfield MR: Peripheral blood antigen-presenting cells from African-Americans exhibit increased CD80 and CD86 expression. Clin Exp Immunol 1999;118:247–252.

27 Shevde LA, Joshi NN, Shinde SR, Nadkarni JJ: Studies on functional status of circulating lymphocytes in unaffected members from cancer families. Hum Immunol 1998;59:373–381.

28 Ganguly A, Rock MJ, Prockop DJ: Conformation-sensitive gel electrophoresis for rapid detection of single-base differences in double-stranded PCR products and DNA fragments: evidence for solvent-induced bends in DNA heteroduplexes [published erratum appears in Proc Natl Acad Sci USA 1994;91:5217]. Proc Natl Acad Sci USA 1993;90:10325–10329.

29 Harper K, Balzano C, Rouvier E, Mattel M, Luciani M, Goldstein P: CTLA4 and CD28 activated lymphocyte moleculaes are closely related in both mouse and human as to sequence, message expression, gene structure and chromosomal location. J Immunol 1991;47:1037–1044.

30 Pociot F, Molvig J, Wogensen L, Worsaae H, Nerup J: A TaqI polymorphism in the human interleukin-1 beta (IL-1 beta) gene correlates with IL-1 beta secretion in vitro. Eur J Clin Invest 1992;22:396–402.

31 Guasch JF, Bertina RM, Reitsma PH: Five novel intragenic dimorphisms in the human interleukin-1 genes combine to high informativity. Cytokine 1996;8:598–602.

32 Denny P, Lord CJ, Hill NJ, Goy JV, Levy ER, Podolin PL, Peterson LB, Wicker LS, Todd JA, Lyons PA: Mapping of the IDDM locus Idd3 to a 0.35-cM interval containing the interleukin-2 gene. Diabetes 1997;46:695–700.

33 Nolsoe RL, Kristiansen OP, Sangthongpitag K, Larsen ZM, Johannesen J, Karlsen AE, Pociot F, Nerup J, Verge CF, Mandrup-Poulsen T: Complete molecular scanning of the human Fas gene: mutational analysis and linkage studies in families with type I diabetes mellitus. The Danish Study Group of Diabetes in Childhood and The Danish IDDM Epidemiology and Genetics Group. Diabetologia 2000;43: 800–808.

34 D'Alfonso S, Richiardi PM: A polymorphic variation in a putative regulation box of the TNFA promoter region. Immunogenetics 1994;39:150–154.

35 Wilson AG, de Vries N, Pociot F, di Giovine FS, van der Putte LB, Duff GW: An allelic polymorphism within the human tumor necrosis factor alpha promoter region is strongly associated with HLA A1, B8, and DR3 alleles. J Exp Med 1993;177:557–560.

36 Tounas NA, Cominelli F: Identification and initial characterization of two polymorphisms in the human Il-10 promoter. Eur Cytokine Network 1996;7:578.

37 Olomolaiye OO, Wood NAP, Bidwell JL: Identification of a novel polymorphism in the IL-6 promoter region. Immunology 1997;92:66.

38 Messer G, Spengler U, Jung MC, Honold G, Eisenburg J, Scholz S, Albert ED, Pape GR, Riethmuller G, Weiss EH: Allelic variation in the TNF-beta gene does not explain the low TNF-beta response in patients with primary biliary cirrhosis. Scand J Immunol 1991;34(6): 735–740.
39 Messer G, Spengler U, Jung MC, Honold G, Blomer K, Pape GR, Riethmuller G, Weiss EH: Polymorphic structure of the tumor necrosis factor (TNF) locus: an NcoI polymorphism in the first intron of the human TNF-beta gene correlates with a variant amino acid in position 26 and a reduced level of TNF-beta production. J Exp Med. 1991;173:209–219.
40 Abraham L, Kroeger K: Impact of the −308 TNF promoter polymorphism on the transcriptional regulation of the TNF gene: relevance to disease. J Leukoc Biol 1999;66:562–566.
41 Viola J, Rao A: Molecular regulation of cytokine gene expression during the immune response. J Clin Immunol 1999;19:98–108.
42 Terry CF, Loukaci V, Green FR: Cooperative influence of genetic polymorphisms on interleukin 6 transcriptional regulation. J Biol Chem 2000;275:18138–18144.
43 Fishman D, Faulds G, Jeffery R, Mohamed-Ali V, Yudkin JS, Humphries S, Woo P: The effect of novel polymorphisms in the interleukin-6 (IL-6) gene on IL-6 transcription and plasma IL-6 levels, and an association with systemic-onset juvenile chronic arthritis. J Clin Invest 1998;102:1369–1376.
44 Morse H, Olomolaiye O, Wood N, Keen L, Bidwell J: Induced heteroduplex genotyping of TNF-alpha, IL-1beta, IL-6 and IL-10 polymorphisms associated with transcriptional regulation. Cytokine 1999;11:789–795.
45 Wilson AG, Symons JA, McDowell TL, McDevitt HO, Duff GW: Effects of a polymorphism in the human tumor necrosis factor α promoter on transcriptional activation. Proc Natl Acad Sci USA 1997;94:3195–3199.
46 Pociot F, Wilson AG, Nerup J, Duff GW: No independent association between a tumor necrosis factor-alpha promotor region polymorphism and insulin-dependent diabetes mellitus. Eur J Immunol 1993;23:3050–3053.
47 Turner DM, Grant SC, Lamb WR, Brenchley PE, Dyer PA, Sinnott PJ, Hutchinson IV: A genetic marker of high TNF-alpha production in heart transplant recipients. Transplantation 1995;60:1113–1117.
48 Sotgiu S, Pugliatti M, Serra C, Rosati G, Dolei A, Marrosu MG: Tumor necrosis factor 2 allele does not contribute to increased tumor necrosis factor-alpha production in Sardinian multiple sclerosis. Ann Neurol 1999;46:799–800.
49 Maurer M, Kruse N, Giess R, Kyriallis K, Toyka KV, Rieckmann P: Gene polymorphism at position −308 of the tumor necrosis factor alpha promotor is not associated with disease progression in multiple sclerosis patients. J Neurol 1999;246:949–954.
50 Wilson AG, Symons JA, McDowell TL, McDevitt HO, Duff GW: Effects of a polymorphism in the human tumor necrosis factor alpha promoter on transcriptional activation. Proc Natl Acad Sci USA 1997;94:3195–3199.
51 McGuire W, Hill AV, Allsopp CE, Greenwood BM, Kwiatkowski D: Variation in the TNF-alpha promoter region associated with susceptibility to cerebral malaria. Nature 1994;371: 508–510.
52 Tian H, Brody LC, Fan S, Huang Z, Landers JP: Capillary and microchip electrophoresis for rapid detection of known mutations by combining allele-specific DNA amplification with heteroduplex analysis. Clin Chem 2001;47: 173–185.
53 Gray IC, Campbell DA, Spurr NK: Single nucleotide polymorphisms as tools in human genetics. Hum Mol Genet 2000;9:2403–2408.
54 Gross E, Arnold N, Goette J, Schwarz-Boeger U, Kiechle M: A comparison of BRCA1 mutation analysis by direct sequencing, SSCP and DHPLC. Hum Genet 1999;105:72–78.
55 Wagner TMU, Stoppa-Lyonnet D, Fleischmann E, Muhr D, Pages S, Sandberg T, Caux V, Moeslinger R, Langbauer G, Borg A, Oefner P: Denaturing high-performance liquid chromatography detects reliably BRCA1 and BRCA2 mutations. Genomics 1999;62:369–376.
56 Phillips J, Smith ED: Breast cancer control and African American women: a review. Cancer Invest 2001;19:273–280.
57 Rose DP, Royak-Schaler R: Tumor biology and prognosis in black breast cancer patients: a review. Cancer Detect Prevent 2001;25:16–31.
58 Shavers VL, Brown ML: Racial and ethnic disparities in the receipt of cancer treatment. J Nat Cancer Inst 2002;94:334–357.
59 Richardus JH, Kunst AE: Black-white differences in infectious disease mortality in the United States. Am J Public Health 2001;91: 1251–1253.
60 Chouchane L, Sfar I, Bousaffara R, El Kamel A, Sfar MT, Ismail A: A repeat polymorphism in interleukin-4 gene is highly associated with specific clinical phenotypes of asthma. Int Arch Allergy Immunol 1999;120:50–55.
61 Davies H, Bignell GR, Cox C, Stephens P, Edkins S, Clegg S, Teague J, Woffendin H, Garnett MJ, Bottomley W, Davis N, Dicks E, et al: Mutations of the BRAF gene in human cancer. Nature 2002;417:949–954.
62 Brose MS, Volpe P, Feldman M, Kumar M, Rishi I, Gerrero R, Einhorn E, Herlyn M, Minna J, Nicholson A, Roth JA, Albelda SM, et al: BRAF and RAS mutations in human lung cancer and melanoma. Cancer Res 2002;62:6997–7000.
63 den Dunnen JT, Antonarakis SE: Nomenclature for the description of human sequence variations. Hum Genet 2001;109:121–124; DOI 10.1007/s004390100505.

```
      log:  D:\Users\1TIM\BRCA1-BRCA2\il6\il6 analyses 6.25.03.log
  log type:  text
 opened on:  25 Jun 2003, 09:57:17

. tab bc gg, all

           |          gg
        bc |         0          1 |     Total
-----------+----------------------+----------
         0 |        94         87 |       181
         1 |       145        100 |       245
-----------+----------------------+----------
     Total |       239        187 |       426

          Pearson chi2(1) =   2.2217   Pr = 0.136
 likelihood-ratio chi2(1) =   2.2199   Pr = 0.136
               Cramer's V =  -0.0722
                    gamma =  -0.1460  ASE = 0.097
          Kendall's tau-b =  -0.0722  ASE = 0.048

. logistic bc gg yob

Logit estimates                                   Number of obs   =        426
                                                  LR chi2(2)      =      70.06
                                                  Prob > chi2     =     0.0000
Log likelihood =  -255.4237                       Pseudo R2       =     0.1206

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .7280581   .1561012    -1.48   0.139     .4782576    1.108333
         yob |    .932738    .009039    -7.19   0.000      .915189    .9506235
------------------------------------------------------------------------------

. sort talc

. by talc: logistic bc gg yob

-------------------------------------------------------------------------------
> _
-> talc = 0

Logit estimates                                   Number of obs   =         98
                                                  LR chi2(2)      =      23.80
                                                  Prob > chi2     =     0.0000
Log likelihood =  -53.05961                       Pseudo R2       =     0.1832

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .3268306   .1594466    -2.29   0.022     .1256198    .8503298
         yob |   .9228287   .0196129    -3.78   0.000     .8851778     .962081
------------------------------------------------------------------------------


-------------------------------------------------------------------------------
> _
-> talc = 1

note: gg~=1 predicts success perfectly
      gg dropped and 6 obs not used
```

```
Logit estimates                                   Number of obs   =         10
                                                  LR chi2(1)      =       0.01
                                                  Prob > chi2     =     0.9324
Log likelihood = -6.7265184                       Pseudo R2       =     0.0005

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         yob |   1.003607   .0426358     0.08   0.932     .9234271    1.090749
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = .

Logit estimates                                   Number of obs   =        312
                                                  LR chi2(2)      =      49.32
                                                  Prob > chi2     =     0.0000
Log likelihood = -181.22884                       Pseudo R2       =     0.1198

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   1.052726   .2717192     0.20   0.842     .6347632    1.745899
         yob |   .9294848   .0111704    -6.08   0.000     .9078469    .9516383
------------------------------------------------------------------------------

. sort qsmoke

. by qsmoke: logistic bc gg yob

-------------------------------------------------------------------------------
> _
-> qsmoke = 0

Logit estimates                                   Number of obs   =        211
                                                  LR chi2(2)      =      34.93
                                                  Prob > chi2     =     0.0000
Log likelihood = -126.50182                       Pseudo R2       =     0.1213

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .7358288   .2243124    -1.01   0.314     .4048496    1.337395
         yob |   .9351053   .0123204    -5.09   0.000     .9112669    .9595672
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = 1

Logit estimates                                   Number of obs   =        146
                                                  LR chi2(2)      =      24.43
                                                  Prob > chi2     =     0.0000
Log likelihood = -88.310294                       Pseudo R2       =     0.1215

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .4264165   .1564026    -2.32   0.020     .2077912    .8750662
         yob |   .9315967   .0169059    -3.90   0.000     .8990441     .965328
------------------------------------------------------------------------------
```

```
> _
-> qsmoke = .

Logit estimates                                   Number of obs   =         69
                                                  LR chi2(2)      =      17.37
                                                  Prob > chi2     =     0.0002
Log likelihood = -36.490659                       Pseudo R2       =     0.1923

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   2.696664   1.613902     1.66   0.097     .8344476    8.714746
         yob |   .9207283   .0224814    -3.38   0.001     .8777033    .9658624
------------------------------------------------------------------------------

. sort hpar

. by hpar: logistic bc gg yob

-------------------------------------------------------------------------------
> _
-> hpar = 0

Logit estimates                                   Number of obs   =        252
                                                  LR chi2(2)      =      40.30
                                                  Prob > chi2     =     0.0000
Log likelihood =  -153.5637                       Pseudo R2       =     0.1160

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .7204084    .197428    -1.20   0.231     .4210241    1.232681
         yob |   .9277755   .0132208    -5.26   0.000     .9022217     .954053
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> hpar = 1

Logit estimates                                   Number of obs   =        174
                                                  LR chi2(2)      =      27.84
                                                  Prob > chi2     =     0.0000
Log likelihood = -101.56929                       Pseudo R2       =     0.1205

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |    .729203   .2564617    -0.90   0.369     .3659988    1.452838
         yob |   .9348939   .0133844    -4.70   0.000     .9090256    .9614983
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> hpar = .
no observations

. sort qdrink

. by qdrink: logistic bc gg yob

-------------------------------------------------------------------------------
```

```
> _
-> qdrink = 1

Logit estimates                                   Number of obs   =         97
                                                  LR chi2(2)      =      13.06
                                                  Prob > chi2     =     0.0015
Log likelihood = -56.304528                       Pseudo R2       =     0.1039

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .4340262   .2002419    -1.81   0.070      .175714    1.072076
         yob |    .948612    .017195    -2.91   0.004     .9155022    .9829194
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qdrink = 2

Logit estimates                                   Number of obs   =        174
                                                  LR chi2(2)      =      42.78
                                                  Prob > chi2     =     0.0000
Log likelihood = -99.204338                       Pseudo R2       =     0.1774


------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .5919745   .2048856    -1.51   0.130     .3003983    1.166564
         yob |   .8960868   .0189681    -5.18   0.000     .8596708    .9340455
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qdrink = .

Logit estimates                                   Number of obs   =        155
                                                  LR chi2(2)      =      18.85
                                                  Prob > chi2     =     0.0001
Log likelihood =  -94.47515                       Pseudo R2       =     0.0907

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   1.291522   .4632566     0.71   0.476     .6394169    2.608673
         yob |   .9441914   .0138095    -3.93   0.000     .9175096    .9716492
------------------------------------------------------------------------------

. logistic ovca gg yob hpar qocuse

Logit estimates                                   Number of obs   =        319
                                                  LR chi2(4)      =      14.21
                                                  Prob > chi2     =     0.0066
Log likelihood = -90.072563                       Pseudo R2       =     0.0731

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |    1.10131   .4427961     0.24   0.810     .5008103    2.421845
         yob |   .9651031   .0156444    -2.19   0.028     .9349227    .9962578
        hpar |    .387173    .189667    -1.94   0.053      .148225     1.01132
      qocuse |   .5044921   .2414534    -1.43   0.153     .1974528    1.288978
------------------------------------------------------------------------------
```

```
. sort talc

. by talc: logistic ovca gg yob hpar qocuse

-------------------------------------------------------------------------------
> _
-> talc = 0

Logit estimates                                   Number of obs   =         97
                                                  LR chi2(4)      =       2.73
                                                  Prob > chi2     =     0.6044
Log likelihood = -23.778934                       Pseudo R2       =     0.0542

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .7525594   .6088031    -0.35   0.725     .1541482    3.674033
         yob |   .9696332   .0335509    -0.89   0.373      .906055    1.037673
        hpar |   .4126426   .3816507    -0.96   0.339     .0673446    2.528397
      qocuse |   .5787886   .5693134    -0.56   0.578     .0841886    3.979117
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = 1

note: hpar~=0 predicts failure perfectly
      hpar dropped and 1 obs not used

Logit estimates                                   Number of obs   =         14
                                                  LR chi2(3)      =       2.64
                                                  Prob > chi2     =     0.4508
Log likelihood = -7.0565902                       Pseudo R2       =     0.1575

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .0944443   .1572231    -1.42   0.156     .0036155    2.467096
         yob |   .9767491   .0717594    -0.32   0.749     .8457603    1.128025
      qocuse |   3.299803   8.422853     0.47   0.640     .0221702    491.1422
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = .

Logit estimates                                   Number of obs   =        207
                                                  LR chi2(4)      =      10.19
                                                  Prob > chi2     =     0.0373
Log likelihood = -56.060241                       Pseudo R2       =     0.0833

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   1.697876   .8749171     1.03   0.304     .6184165    4.661556
         yob |    .959876   .0207953    -1.89   0.059     .9199712    1.001512
        hpar |   .3988716   .2589827    -1.42   0.157     .1117269    1.423994
      qocuse |   .4394548   .2532782    -1.43   0.154     .1420127    1.359882
------------------------------------------------------------------------------

. sort qsmoke
```

```
. by qsmoke: logistic ovca gg yob hpar qocuse

-------------------------------------------------------------------------------
> _
-> qsmoke = 0

Logit estimates                                   Number of obs   =        182
                                                  LR chi2(4)      =      11.29
                                                  Prob > chi2     =     0.0235
Log likelihood = -48.532897                       Pseudo R2       =     0.1042

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .7843681   .4324235    -0.44   0.660     .2662263    2.310941
         yob |   .9773061   .0199579    -1.12   0.261     .9389617    1.017216
        hpar |   .4590244   .2882035    -1.24   0.215      .134091    1.571347
      qocuse |   .2575592   .1687983    -2.07   0.038     .0712885    .9305394
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = 1

Logit estimates                                   Number of obs   =        133
                                                  LR chi2(4)      =       5.86
                                                  Prob > chi2     =     0.2100
Log likelihood = -39.644106                       Pseudo R2       =     0.0688

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   1.459844   .8995363     0.61   0.539     .4363226    4.884334
         yob |   .9542176   .0254588    -1.76   0.079     .9056015    1.005444
        hpar |   .2719496   .2288206    -1.55   0.122     .0522734    1.414804
      qocuse |   1.080734   .9064691     0.09   0.926     .2088194    5.593284
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = .
outcome does not vary; remember:
                                0 = negative outcome,
        all other nonmissing values = positive outcome

. sort hpar

. by hpar: logistic ovca gg yob qocuse

-------------------------------------------------------------------------------
> _
-> hpar = 0

Logit estimates                                   Number of obs   =        212
                                                  LR chi2(3)      =       8.58
                                                  Prob > chi2     =     0.0354
Log likelihood = -66.366569                       Pseudo R2       =     0.0607

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .7507918   .3522135    -0.61   0.541     .2993653    1.882945
```

```
        yob |   .9586045   .0192244    -2.11   0.035     .9216563    .9970339
      qocuse |   .7619607   .4571118    -0.45   0.650     .2351169    2.469342
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> hpar = 1

Logit estimates                                   Number of obs   =        107
                                                  LR chi2(3)      =       8.56
                                                  Prob > chi2     =     0.0357
Log likelihood = -21.574083                       Pseudo R2       =     0.1656

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   3.624951   3.219578     1.45   0.147     .6357543    20.66879
         yob |   .9828544   .0327149    -0.52   0.603     .9207813    1.049112
      qocuse |   .1761236   .1716203    -1.78   0.075     .0260844      1.1892
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> hpar = .
no observations

. sort qocuse

. sort qdrink

. by qdrink: logistic ovca gg yob qocuse

-------------------------------------------------------------------------------
> _
-> qdrink = 1

Logit estimates                                   Number of obs   =         96
                                                  LR chi2(3)      =       2.71
                                                  Prob > chi2     =     0.4384
Log likelihood = -32.819925                       Pseudo R2       =     0.0397

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   1.047816   .6870608     0.07   0.943     .2898318    3.788124
         yob |   .9611939   .0262369    -1.45   0.147     .9111219    1.014018
      qocuse |   3.127558   2.776885     1.28   0.199     .5488372    17.82244
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qdrink = 2

Logit estimates                                   Number of obs   =        172
                                                  LR chi2(3)      =      12.62
                                                  Prob > chi2     =     0.0055
Log likelihood = -39.760889                       Pseudo R2       =     0.1369

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .8385978   .5244535    -0.28   0.778     .2461602    2.856864
         yob |   .9603528   .0241293    -1.61   0.107     .9142058    1.008829
```

```
      qocuse |   .2185881   .1464132    -2.27   0.023      .0588139      .812405
-------------------------------------------------------------------------------


-> qdrink = .

Logit estimates                                   Number of obs   =         51
                                                  LR chi2(3)      =       4.80
                                                  Prob > chi2     =     0.1872
Log likelihood = -13.959386                       Pseudo R2       =     0.1467

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   1.682076   1.730298     0.51   0.613     .2239969    12.63134
         yob |    1.00282   .0411655     0.07   0.945     .9252971    1.086837
      qocuse |   .1150245   .1253233    -1.98   0.047     .0135949     .973205
------------------------------------------------------------------------------

. sort qocuse

. by qocuse: logistic ovca gg yob hpar


-> qocuse = 1

Logit estimates                                   Number of obs   =         75
                                                  LR chi2(3)      =       0.40
                                                  Prob > chi2     =     0.9403
Log likelihood = -34.385124                       Pseudo R2       =     0.0058

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .7200264   .4519709    -0.52   0.601     .2103958    2.464109
         yob |   .9961279   .0187946    -0.21   0.837      .959964    1.033654
        hpar |    .833589    .541899    -0.28   0.779     .2331328    2.980578
------------------------------------------------------------------------------


-> qocuse = 2

Logit estimates                                   Number of obs   =        244
                                                  LR chi2(3)      =      17.20
                                                  Prob > chi2     =     0.0006
Log likelihood = -50.458821                       Pseudo R2       =     0.1456

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   1.307191   .7135752     0.49   0.624      .448416    3.81063
         yob |   .8943634   .0297191    -3.36   0.001     .8379713    .9545505
        hpar |   .0783524   .0755985    -2.64   0.008     .0118241    .5192012
------------------------------------------------------------------------------


-> qocuse = .
```

```
Logit estimates                                   Number of obs   =        104
                                                  LR chi2(3)      =      10.11
                                                  Prob > chi2     =     0.0177
Log likelihood = -41.264445                       Pseudo R2       =     0.1091

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .6787254   .4131013    -0.64   0.524     .2058811    2.237545
         yob |    .943631   .0190277    -2.88   0.004     .9070647    .9816715
        hpar |   1.772137    1.07995     0.94   0.348     .5367477    5.850922
------------------------------------------------------------------------------

. by qocuse: logistic bc gg yob hpar

-------------------------------------------------------------------------------
> _
-> qocuse = 1

Logit estimates                                   Number of obs   =         75
                                                  LR chi2(3)      =      13.19
                                                  Prob > chi2     =     0.0042
Log likelihood = -41.143346                       Pseudo R2       =     0.1382

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .5643434   .3030322    -1.07   0.287     .1970057    1.616621
         yob |   .9443994   .0177095    -3.05   0.002     .9103195    .9797551
        hpar |   .4395179   .2578881    -1.40   0.161     .1391657    1.388101
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qocuse = 2

Logit estimates                                   Number of obs   =        244
                                                  LR chi2(3)      =      46.53
                                                  Prob > chi2     =     0.0000
Log likelihood = -145.46015                       Pseudo R2       =     0.1379

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   .6732572   .1919584    -1.39   0.165     .3850236    1.177266
         yob |    .903982   .0162841    -5.60   0.000     .8726226    .9364683
        hpar |   1.017104   .3280572     0.05   0.958     .5405286     1.91387
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qocuse = .

Logit estimates                                   Number of obs   =        107
                                                  LR chi2(3)      =      13.98
                                                  Prob > chi2     =     0.0029
Log likelihood = -64.227135                       Pseudo R2       =     0.0982

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
          gg |   1.047061   .4671832     0.10   0.918     .4366955    2.510531
         yob |   .9439805   .0162587    -3.35   0.001     .9126459     .976391
```

```
        hpar |   .7520058   .3320499    -0.65   0.519      .3164981    1.786781
------------------------------------------------------------------------------

. tab il6 gg

  IL6/-174 |          gg
      type |         0          1 |     Total
-----------+----------------------+----------
         1 |         0        225 |       225
         2 |       220          0 |       220
         3 |        67          0 |        67
-----------+----------------------+----------
     Total |       287        225 |       512


. tab anyg

      anyg |      Freq.     Percent        Cum.
-----------+-----------------------------------
         0 |         67       13.09       13.09
         1 |        445       86.91      100.00
-----------+-----------------------------------
     Total |        512      100.00

. tab il6 anyg

  IL6/-174 |          anyg
      type |         0          1 |     Total
-----------+----------------------+----------
         1 |         0        225 |       225
         2 |         0        220 |       220
         3 |        67          0 |        67
-----------+----------------------+----------
     Total |        67        445 |       512


. logistic bc anyg yob

Logit estimates                                   Number of obs   =        426
                                                  LR chi2(2)      =      67.88
                                                  Prob > chi2     =     0.0000
Log likelihood = -256.51327                       Pseudo R2       =     0.1169

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |    .957859   .3080327    -0.13   0.893      .5099978    1.799015
         yob |   .9329447   .0090132    -7.18   0.000      .9154453    .9507786
------------------------------------------------------------------------------

. logistic ovca anyg hpar qocuse yob

Logit estimates                                   Number of obs   =        319
                                                  LR chi2(4)      =      15.10
                                                  Prob > chi2     =     0.0045
Log likelihood = -89.627943                       Pseudo R2       =     0.0777

------------------------------------------------------------------------------
         ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .5410087   .3252652    -1.02   0.307      .1665118    1.757776
        hpar |   .3791563   .1858294    -1.98   0.048      .1450886    .9908393
      qocuse |   .5238514   .2540874    -1.33   0.183      .2024621    1.355415
```

```
         yob |   .9636326   .0159071    -2.24   0.025      .9329541    .9953198
------------------------------------------------------------------------------

. sort talc

. by talc: logistic bc anyg yob

-------------------------------------------------------------------------------
> _
-> talc = 0

Logit estimates                                   Number of obs   =         98
                                                  LR chi2(2)      =      18.35
                                                  Prob > chi2     =     0.0001
Log likelihood =  -55.78559                       Pseudo R2       =     0.1412

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .7619756   .5821626    -0.36   0.722      .1704544    3.406229
         yob |   .9245411   .0194918    -3.72   0.000      .8871164    .9635446
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = 1

note: anyg~=1 predicts success perfectly
      anyg dropped and 1 obs not used

Logit estimates                                   Number of obs   =         15
                                                  LR chi2(1)      =       0.02
                                                  Prob > chi2     =     0.8903
Log likelihood = -8.6892242                       Pseudo R2       =     0.0011

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         yob |   .9949251   .0367043    -0.14   0.890      .9255251    1.069529
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = .

Logit estimates                                   Number of obs   =        312
                                                  LR chi2(2)      =      49.32
                                                  Prob > chi2     =     0.0000
Log likelihood = -181.23106                       Pseudo R2       =     0.1198

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   1.073373   .4044781     0.19   0.851      .5128536    2.246509
         yob |   .9293647   .0111927    -6.08   0.000      .9076843    .9515629
------------------------------------------------------------------------------

. by talc: logistic ovca anyg yob

-------------------------------------------------------------------------------
> _
-> talc = 0
```

```
Logit estimates                                   Number of obs   =         98
                                                  LR chi2(2)      =       1.70
                                                  Prob > chi2     =     0.4272
Log likelihood = -24.366631                       Pseudo R2       =     0.0337

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .4452307   .5352909    -0.67   0.501     .0421895    4.698569
         yob |     .96338   .0281899    -1.27   0.202     .9096834    1.020246
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = 1

note: anyg~=1 predicts success perfectly
      anyg dropped and 1 obs not used

Logit estimates                                   Number of obs   =         14
                                                  LR chi2(1)      =       0.00
                                                  Prob > chi2     =     0.9772
Log likelihood = -7.2737101                       Pseudo R2       =     0.0001

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         yob |   1.001142    .040005     0.03   0.977     .9257253    1.082702
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = .

Logit estimates                                   Number of obs   =        310
                                                  LR chi2(2)      =      17.72
                                                  Prob > chi2     =     0.0001
Log likelihood = -100.42788                       Pseudo R2       =     0.0811

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .4164998    .184591    -1.98   0.048     .1747293    .9928047
         yob |   .9540375   .0122472    -3.67   0.000     .9303329    .9783461
------------------------------------------------------------------------------

. sort qsmoke

. by qsmoke: logistic bc anyg yob

-------------------------------------------------------------------------------
> _
-> qsmoke = 0

Logit estimates                                   Number of obs   =        211
                                                  LR chi2(2)      =      34.52
                                                  Prob > chi2     =     0.0000
Log likelihood = -126.70847                       Pseudo R2       =     0.1199

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
        anyg |   1.394707   .5982843     0.78   0.438     .6016516    3.233114
         yob |   .9354804   .0122899    -5.08   0.000     .9117002    .9598809
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = 1

Logit estimates                                   Number of obs   =        146
                                                  LR chi2(2)      =      19.71
                                                  Prob > chi2     =     0.0001
Log likelihood =  -90.67327                       Pseudo R2       =     0.0980

-------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------
        anyg |   .5193947   .3849374    -0.88   0.377     .1215203    2.219964
         yob |   .9342617   .0165928    -3.83   0.000        .9023    .9673556
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = .

Logit estimates                                   Number of obs   =         69
                                                  LR chi2(2)      =      14.63
                                                  Prob > chi2     =     0.0007
Log likelihood = -37.863363                       Pseudo R2       =     0.1619

-------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------
        anyg |   .7501741   .5469725    -0.39   0.693     .1796916    3.131817
         yob |   .9252658   .0222491    -3.23   0.001     .8826699    .9699173
-------------------------------------------------------------------------------

. by qsmoke: logistic ovca anyg yob

-------------------------------------------------------------------------------
> _
-> qsmoke = 0

Logit estimates                                   Number of obs   =        210
                                                  LR chi2(2)      =      10.48
                                                  Prob > chi2     =     0.0053
Log likelihood = -63.029767                       Pseudo R2       =     0.0767

-------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+-----------------------------------------------------------------
        anyg |     .61457   .3776313    -0.79   0.428     .1843056    2.049294
         yob |    .952772   .0145772    -3.16   0.002     .9246255    .9817754
-------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = 1

Logit estimates                                   Number of obs   =        146
                                                  LR chi2(2)      =       5.86
                                                  Prob > chi2     =     0.0533
Log likelihood =   -40.9137                       Pseudo R2       =     0.0669
```

```
------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .1833256   .1417593    -2.19   0.028     .0402734    .8345025
         yob |   .9689036    .022045    -1.39   0.165     .9266454    1.013089
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = .

Logit estimates                                   Number of obs   =         67
                                                  LR chi2(2)      =       6.32
                                                  Prob > chi2     =     0.0425
Log likelihood = -28.334136                       Pseudo R2       =     0.1003

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .3994184   .2986292    -1.23   0.220     .0922611    1.729169
         yob |   .9518432   .0225517    -2.08   0.037     .9086532    .9970862
------------------------------------------------------------------------------

. sort hpar

. by hpar: logistic bc anyg yob

-------------------------------------------------------------------------------
> _
-> hpar = 0

Logit estimates                                   Number of obs   =        252
                                                  LR chi2(2)      =      39.20
                                                  Prob > chi2     =     0.0000
Log likelihood = -154.11285                       Pseudo R2       =     0.1128

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .7734556   .3437497    -0.58   0.563     .3236926    1.848153
         yob |   .9270561   .0131991    -5.32   0.000      .901544    .9532902
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> hpar = 1

Logit estimates                                   Number of obs   =        174
                                                  LR chi2(2)      =      27.16
                                                  Prob > chi2     =     0.0000
Log likelihood = -101.90606                       Pseudo R2       =     0.1176

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   1.188286   .5537645     0.37   0.711     .4766995    2.962085
         yob |    .936525   .0131538    -4.67   0.000     .9110957     .962664
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> hpar = .
```

```
no observations

. by hpar: logistic ovca anyg yob qocuse

-------------------------------------------------------------------------------
> _
-> hpar = 0

Logit estimates                                   Number of obs   =        212
                                                  LR chi2(3)      =       9.02
                                                  Prob > chi2     =     0.0290
Log likelihood = -66.148858                       Pseudo R2       =     0.0638

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .5164408    .359563    -0.95   0.343      .131943    2.021412
         yob |    .956832   .0193447    -2.18   0.029     .9196585     .995508
      qocuse |   .7746706   .4682933    -0.42   0.673     .2368974    2.533225
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> hpar = 1

Logit estimates                                   Number of obs   =        107
                                                  LR chi2(3)      =       6.33
                                                  Prob > chi2     =     0.0968
Log likelihood =  -22.69163                       Pseudo R2       =     0.1223

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .6796332   .8137271    -0.32   0.747     .0650321    7.102661
         yob |   .9789658   .0307875    -0.68   0.499     .9204455    1.041207
      qocuse |   .1908755   .1859174    -1.70   0.089     .0282917    1.287779
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> hpar = .
no observations

. by qsmoke: logistic ovca anyg yob hpar qocuse
not sorted
r(5);

. sort qsmoke

. by qsmoke: logistic ovca anyg yob hpar qocuse

-------------------------------------------------------------------------------
> _
-> qsmoke = 0

Logit estimates                                   Number of obs   =        182
                                                  LR chi2(4)      =      11.31
                                                  Prob > chi2     =     0.0233
Log likelihood = -48.522179                       Pseudo R2       =     0.1044

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
```

```
-------------+----------------------------------------------------------------
        anyg |   1.616739   1.763206     0.44   0.660      .1906883    13.70742
         yob |   .9774389   .0198781    -1.12   0.262      .9392448    1.017186
        hpar |   .4691719   .2957269    -1.20   0.230      .1363972    1.613833
      qocuse |   .2621143   .1701376    -2.06   0.039      .0734477    .9354125
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = 1

Logit estimates                                   Number of obs   =        133
                                                  LR chi2(4)      =       9.00
                                                  Prob > chi2     =     0.0610
Log likelihood = -38.071797                       Pseudo R2       =     0.1057

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .1948151   .1583589    -2.01   0.044      .0396012    .9583783
         yob |   .9499993   .0266861    -1.83   0.068      .8991093     1.00377
        hpar |   .2692116   .2258783    -1.56   0.118      .0519882    1.394064
      qocuse |   1.435402   1.234235     0.42   0.674      .2661126    7.742509
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qsmoke = .
outcome does not vary; remember:
                                0 = negative outcome,
        all other nonmissing values = positive outcome

. by talc: logistic ovca anyg yob hpar qocuse
not sorted
r(5);

. sort talc

. by talc: logistic ovca anyg yob hpar qocuse

-------------------------------------------------------------------------------
> _
-> talc = 0

Logit estimates                                   Number of obs   =         97
                                                  LR chi2(4)      =       3.35
                                                  Prob > chi2     =     0.5005
Log likelihood = -23.465965                       Pseudo R2       =     0.0667

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .3100177   .3864726    -0.94   0.348      .0269318    3.568674
         yob |    .963246   .0350777    -1.03   0.304      .8968911     1.03451
        hpar |   .3709245    .349391    -1.05   0.292      .0585463    2.350021
      qocuse |   .5362799   .5404717    -0.62   0.536      .0743938    3.865863
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = 1

note: anyg~=1 predicts success perfectly
```

```
      anyg dropped and 1 obs not used

note: hpar~=0 predicts failure perfectly
      hpar dropped and 1 obs not used


Logit estimates                                   Number of obs   =         13
                                                  LR chi2(2)      =       0.06
                                                  Prob > chi2     =     0.9690
Log likelihood = -6.9912009                       Pseudo R2       =     0.0045

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         yob |   .9914038   .0621782    -0.14   0.891     .8767293    1.121077
      qocuse |   1.684549    3.67448     0.24   0.811       .02343    121.1144
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> talc = .

Logit estimates                                   Number of obs   =        207
                                                  LR chi2(4)      =       9.14
                                                  Prob > chi2     =     0.0577
Log likelihood = -56.585393                       Pseudo R2       =     0.0747

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .8977047   .7320254    -0.13   0.895     .1815636    4.438521
         yob |   .9618005   .0206529    -1.81   0.070     .9221616    1.003143
        hpar |   .4052332   .2593316    -1.41   0.158     .1156041    1.420486
      qocuse |   .4345278   .2535003    -1.43   0.153     .1384946    1.363334
------------------------------------------------------------------------------

. sort qdrink

. by qdrink: logistic ovca anyg yob hpar qocuse

-------------------------------------------------------------------------------
> _
-> qdrink = 1

note: anyg~=1 predicts failure perfectly
      anyg dropped and 13 obs not used


Logit estimates                                   Number of obs   =         83
                                                  LR chi2(3)      =       2.58
                                                  Prob > chi2     =     0.4602
Log likelihood = -31.174786                       Pseudo R2       =     0.0398

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         yob |   .9612891   .0294887    -1.29   0.198     .9051955    1.020859
        hpar |   .5667915    .422311    -0.76   0.446     .1315834    2.441437
      qocuse |   3.102458   2.821143     1.25   0.213     .5220197    18.43847
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
```

```
> _
-> qdrink = 2

Logit estimates                                   Number of obs   =        172
                                                  LR chi2(4)      =      21.38
                                                  Prob > chi2     =     0.0003
Log likelihood =  -35.37867                       Pseudo R2       =     0.2320

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .1195427   .1006885    -2.52   0.012     .0229389    .6229795
         yob |   .9418447   .0282665    -2.00   0.046     .8880414    .9989078
        hpar |   .2139947   .1877825    -1.76   0.079     .0383235    1.194926
      qocuse |   .3461361   .2630285    -1.40   0.163     .0780584    1.534879
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qdrink = .

note: anyg~=1 predicts failure perfectly
      anyg dropped and 4 obs not used


Logit estimates                                   Number of obs   =         47
                                                  LR chi2(3)      =       4.55
                                                  Prob > chi2     =     0.2083
Log likelihood = -13.654985                       Pseudo R2       =     0.1427

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         yob |   .9949205   .0426545    -0.12   0.905     .9147353    1.082135
        hpar |   .4183618   .5641847    -0.65   0.518     .0297617    5.880927
      qocuse |   .1163993   .1293496    -1.94   0.053      .013184    1.027671
------------------------------------------------------------------------------

.
. by qdrink: logistic bc anyg yob

-------------------------------------------------------------------------------
> _
-> qdrink = 1

Logit estimates                                   Number of obs   =         97
                                                  LR chi2(2)      =      10.02
                                                  Prob > chi2     =     0.0067
Log likelihood = -57.823733                       Pseudo R2       =     0.0797

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .6878239     .45767    -0.56   0.574       .18668    2.534293
         yob |   .9492039   .0171656    -2.88   0.004     .9161492    .9834511
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qdrink = 2

Logit estimates                                   Number of obs   =        174
```

```
                                                LR chi2(2)      =      40.73
                                                Prob > chi2     =     0.0000
Log likelihood = -100.23093                     Pseudo R2       =     0.1689

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .7234206    .455694    -0.51   0.607     .2104775     2.48643
         yob |   .8961403   .0189823    -5.18   0.000     .8596975      .934128
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qdrink = .

Logit estimates                                 Number of obs   =        155
                                                LR chi2(2)      =      19.01
                                                Prob > chi2     =     0.0001
Log likelihood = -94.395298                     Pseudo R2       =     0.0915

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   1.472146   .6931306     0.82   0.411     .5850321     3.704437
         yob |   .9430202    .013992    -3.95   0.000     .9159913     .9708466
------------------------------------------------------------------------------

.
. sort qocuse

. by qocuse: logistic bc anyg yob

-------------------------------------------------------------------------------
> _
-> qocuse = 1

Logit estimates                                 Number of obs   =         75
                                                LR chi2(2)      =      10.05
                                                Prob > chi2     =     0.0066
Log likelihood = -42.711577                     Pseudo R2       =     0.1053

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   1.273363   1.028914     0.30   0.765     .2613101     6.205088
         yob |   .9522282    .015928    -2.93   0.003     .9215162     .9839638
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qocuse = 2

Logit estimates                                 Number of obs   =        244
                                                LR chi2(2)      =      45.41
                                                Prob > chi2     =     0.0000
Log likelihood = -146.02175                     Pseudo R2       =     0.1346

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .6361947   .3117478    -0.92   0.356      .243492     1.662246
         yob |   .9025507   .0160651    -5.76   0.000     .8716066     .9345935
------------------------------------------------------------------------------
```

```
> _
-> qocuse = .

Logit estimates                                   Number of obs   =        107
                                                  LR chi2(2)      =      13.67
                                                  Prob > chi2     =     0.0011
Log likelihood = -64.384128                       Pseudo R2       =     0.0960

------------------------------------------------------------------------------
          bc | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   1.192324   .6152349     0.34   0.733     .4336876    3.278017
         yob |   .9438679   .0162313    -3.36   0.001     .9125852     .976223
------------------------------------------------------------------------------

. by qocuse: logistic ovca anyg yob hpar

-------------------------------------------------------------------------------
> _
-> qocuse = 1

Logit estimates                                   Number of obs   =         75
                                                  LR chi2(3)      =       2.33
                                                  Prob > chi2     =     0.5073
Log likelihood = -33.421379                       Pseudo R2       =     0.0336

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .2824761   .2297036    -1.55   0.120     .0573856    1.390467
         yob |   .9939453   .0192227    -0.31   0.754     .9569747    1.032344
        hpar |   .7647594   .5063319    -0.41   0.685      .208911    2.799551
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qocuse = 2

Logit estimates                                   Number of obs   =        244
                                                  LR chi2(3)      =      16.96
                                                  Prob > chi2     =     0.0007
Log likelihood = -50.577541                       Pseudo R2       =     0.1436

------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   1.079586   1.207083     0.07   0.945     .1206509     9.66015
         yob |   .8942777   .0295537    -3.38   0.001     .8381896    .9541191
        hpar |   .0793689   .0757055    -2.66   0.008     .0122388    .5147105
------------------------------------------------------------------------------

-------------------------------------------------------------------------------
> _
-> qocuse = .

Logit estimates                                   Number of obs   =        104
                                                  LR chi2(3)      =      12.67
                                                  Prob > chi2     =     0.0054
Log likelihood =  -39.98362                       Pseudo R2       =     0.1368
```

```
------------------------------------------------------------------------------
        ovca | Odds Ratio   Std. Err.      z    P>|z|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
        anyg |   .3476912   .2089978    -1.76   0.079     .1070373    1.129412
         yob |   .9454642   .0192715    -2.75   0.006     .9084373    .9840002
        hpar |   1.686869   1.037206     0.85   0.395     .5054787    5.629372
------------------------------------------------------------------------------

. sum aad

    Variable |     Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         aad |    1013    41.31589    9.980263        21         77

. kwallis aad, by(gg)

Test: Equality of populations (Kruskal-Wallis test)

       gg          _Obs    _RankSum
        0           142    17227.00
        1            99    11934.00

chi-squared =     0.007 with 1 d.f.
probability =     0.9326

chi-squared with ties =     0.007 with 1 d.f.
probability =     0.9326

. kwallis aad, by(anyg)

Test: Equality of populations (Kruskal-Wallis test)

     anyg          _Obs    _RankSum
        0           32     3962.00
        1          209    25199.00

chi-squared =     0.060 with 1 d.f.
probability =     0.8064

chi-squared with ties =     0.060 with 1 d.f.
probability =     0.8062

. kwallis qwgt18, by(anyg)

Test: Equality of populations (Kruskal-Wallis test)

     anyg          _Obs    _RankSum
        0           27     4069.00
        1          221    26807.00

chi-squared =     4.043 with 1 d.f.
probability =     0.0444

chi-squared with ties =     4.067 with 1 d.f.
probability =     0.0437

. kwallis qwgt30, by(anyg)

Test: Equality of populations (Kruskal-Wallis test)

     anyg          _Obs    _RankSum
        0           18     1987.50
        1          143    11053.50
```

```
chi-squared =     8.068 with 1 d.f.
probability =     0.0045

chi-squared with ties =     8.109 with 1 d.f.
probability =     0.0044

. kwallis qwgt40, by(anyg)

Test: Equality of populations (Kruskal-Wallis test)

     anyg          _Obs   _RankSum
        0            18    1258.50
        1           151   13106.50

chi-squared =     1.914 with 1 d.f.
probability =     0.1665

chi-squared with ties =     1.998 with 1 d.f.
probability =     0.1575

. kwallis qwgtmax, by(anyg)

Test: Equality of populations (Kruskal-Wallis test)

     anyg          _Obs   _RankSum
        0            17    1662.50
        1           141   10898.50

chi-squared =     3.045 with 1 d.f.
probability =     0.0810

chi-squared with ties =     3.049 with 1 d.f.
probability =     0.0808

. summ  qwgt18

    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      qwgt18 |    1060    122.4887    20.2279         0        340

. summ  qwgt30

    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      qwgt30 |     557    125.6302   38.29528         0        425

. sum  qwgt40

    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+--------------------------------------------------------
      qwgt40 |     567    103.4145   65.50591         0        280

. sum qheight

    Variable |     Obs        Mean   Std. Dev.       Min        Max
-------------+--------------------------------------------------------
     qheight |    1101    64.47207   2.914962        49         90

. gen bmi=qwgt18/qheight
(1162 missing values generated)

. replace bmi=bmi/qheight
```

```
(1006 real changes made)

. sum bmi

    Variable |     Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         bmi |    1007    .0296026    .0050195          0    .0884495

. replace bmi=bmi*100
(1006 real changes made)

. sum bmi

    Variable |     Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         bmi |    1007    2.960255    .5019517          0    8.844953

. replace bmi=bmi*10
(1006 real changes made)

. sum bmi

    Variable |     Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         bmi |    1007    29.60255    5.019517          0    88.44952

. replace bmi=. if bmi==0
(1 real change made, 1 to missing)

. sum bmi

    Variable |     Obs        Mean    Std. Dev.       Min        Max
-------------+--------------------------------------------------------
         bmi |    1006    29.63198    4.934349   16.66667    88.44952

. kwallis bmi, by(gg)

Test: Equality of populations (Kruskal-Wallis test)

       gg          _Obs   _RankSum
        0           130   15402.50
        1           114   14487.50

chi-squared =     0.902 with 1 d.f.
probability =     0.3422

chi-squared with ties =     0.902 with 1 d.f.
probability =     0.3421

. kwallis bmi, by(anyg)

Test: Equality of populations (Kruskal-Wallis test)

     anyg          _Obs   _RankSum
        0           26    3624.50
        1          218   26265.50

chi-squared =     1.669 with 1 d.f.
probability =     0.1964

chi-squared with ties =     1.669 with 1 d.f.
probability =     0.1963
```

Table 1. Effects of *IL6-GG* Genotypes

Cancer	Group	OR Associated with GG Genotype (95% CI)
Breast Cancer*	Total Sample	0.73 (0.48-1.11)
	Never Smokers	0.74 (0.40-1.34)
	Ever Smokers	0.43 (0.21-0.88)
	Low Parity (≤2 Live Birth)	0.72 (0.42-1.23)
	High Parity (≥3 Live Births)	0.73 (0.37-1.45)
	Never Oral Contraceptive User	0.56 (0.20-1.62)
	Ever Oral Contraceptive User	0.67 (0.39-1.12)
Ovarian Cancer**	Total Sample	1.10 (0.50-2.42)
	Never Smokers	0.78 (0.27-2.31)
	Ever Smokers	1.46 (0.44-4.88)
	Talc Non-Users	0.75 (0.15-3.67)
	Talc Users	0.09 (0.004-2.47)
	Low Parity (≤2 Live Birth)	0.75 (0.30-1.88)
	High Parity (≥3 Live Births)	3.62 (0.63-20.67)
	Never Oral Contraceptive User	0.72 (0.21-2.46)
	Ever Oral Contraceptive User	1.31 (0.45-3.81)

*OR adjusted for birth year

** OR adjusted for year of birth, parity, and oral contraceptive use (except where factor is the main effect of interest)

Table 2. Effects of *IL6-Any G* Genotypes

Cancer	Group	OR Associated with Any G Genotype (95% CI)
Breast Cancer*	Total Sample	0.96 (0.51-1.80)
	Never Smokers	1.39 (0.60-3.23)
	Ever Smokers	0.52 (0.12-2.22)
	Low Parity (≤2 Live Birth)	0.77 (0.32-1.85)
	High Parity (≥3 Live Births)	1.19 (0.48-2.96)
	Never Oral Contraceptive User	1.27 (0.26-6.21)
	Ever Oral Contraceptive User	0.64 (0.24-1.66)
Ovarian Cancer**	Total Sample	0.54 (0.17-1.76)
	Never Smokers	1.62 (0.19-13.71)
	Ever Smokers	0.19 (0.04-0.96)
	Talc Non-Users	0.31 (0.03-3.57)
	Talc Users	0.99 (0.87-1.12)
	Low Parity (≤2 Live Birth)	0.52 (0.13-2.02)
	High Parity (≥3 Live Births)	0.68 (0.07-7.10)
	Never Oral Contraceptive User	0.28 (0.06-1.39)
	Ever Oral Contraceptive User	1.08 (0.12-9.66)

*OR adjusted for birth year

** OR adjusted for year of birth, parity, and oral contraceptive use (except where factor is the main effect of interest)